Towards Usable Electroencephalography-based Brain-Computer Interfaces

Fabien Lotte

▶ To cite this version:

Fabien Lotte. Towards Usable Electroencephalography-based Brain-Computer Interfaces. Computer Science [cs]. Univ. Bordeaux, 2016. tel-01416980

HAL Id: tel-01416980
https://inria.hal.science/tel-01416980

Submitted on 15 Dec 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

HABILITATION THESIS

UNIVERSITY OF BORDEAUX

DOCTORAL SCHOOL MATHEMATICS AND COMPUTER SCIENCE

by **Fabien LOTTE**

SPECIALTY: COMPUTER SCIENCE

Towards Usable Electroencephalography-based Brain-Computer Interfaces

---

Defended: September 21st, 2016

Habilitation committee:

Gernot MULLER-PUTZ	Professor, TU Graz, Austria	Rapporteur
Benjamin BLANKERTZ	Professor, TU Berlin, Germany	Rapporteur
Frédéric DEHAIS	Professor, ISAE Toulouse, France	Rapporteur
Maureen CLERC	Research Director, Inria Sophia-Antipolis, France	Examinatrice
Laurence NIGAY	Professor, Grenoble Alpes University, France	Examinatrice
Frédéric ALEXANDRE	Research Director, Inria Bordeaux Sud-Ouest	Examinateur

– 2016 –

Science is like love: it is sometimes overwhelmingly complex and regularly crushes your feelings... but damn I enjoy it so much!

Résumé

Title Towards Usable Electroencephalography-based Brain-Computer Interfaces

Abstract Brain-Computer Interfaces (BCIs) are systems that can translate brain activity patterns of a user into messages or commands for an interactive application. Such brain activity is typically measured using Electroencephalography (EEG), before being processed and classified by the system. EEG-based BCIs have proven promising for a wide range of applications ranging from communication and control for motor impaired users, to gaming targeted at the general public, real-time mental state monitoring and stroke rehabilitation, to name a few. Despite this promising potential, BCIs are still scarcely used outside laboratories for practical applications. The main reason preventing EEG-based BCIs from being widely used is arguably their poor usability, which is notably due to their low robustness and reliability, as well as their long calibration and training times. The research presented in this manuscript aims at addressing these different points in order to make EEG-based BCIs usable, i.e., to increase their efficacy and efficiency. In particular, we present a set of contributions towards this goal 1) at the EEG signal processing and classification level, to robustly decode EEG signals and translate them into commands, 2) at the user training level, to ensure that users can learn to control a BCI efficiently and effectively, and 3) at the usage level, to explore novel applications of BCIs for which the current reliability can already be useful.

First, in terms of EEG signal processing tools, we proposed a number of methods to improve BCI reliability, despite EEG signal variability, poor signal-to-noise ratio and high sensitivity to artifacts, as well as to reduce BCIs calibration times. More precisely, we complemented traditionally used features by exploring alternative representations of EEG signals. We also explored and designed regularized spatial filters to learn more robust and stable features. Finally, we propose algorithms to reduce BCIs calibration times, i.e., to calibrate BCIs with as few examples of EEG signals from the target user as possible, by re-using data from previous users or by generating artificial EEG signals. Altogether, these methods enabled an increased BCI classification accuracy, i.e., efficacy, and BCI calibration with much less data than standard approaches do, thus improving their efficiency.

Second, rather than improving EEG signal processing alone, we advocate that BCIs can also be made more usable by guiding users to efficiently learn BCI control mastery. Indeed, BCI control is known to be a skill that needs to be learnt. A study of models and guidelines from educational sciences enabled us to identify many theoretical limitations of current standard BCI training approaches, thus highlighting the need for alternative ones. In particular, educational sciences recommend to train people with adapted and adaptive

training tasks, using explanatory feedback in motivating environments. In contrast standard BCI training protocols are commonly fixed, repetitive, rather boring and provide purely corrective feedback. To address these limitations, we studied what kind of users manage to use a BCI and why. We also explored new feedback types, in particular richer feedback, multi-user feedback and tactile feedback to help users to learn BCI control skills more efficiently. Overall, our studies identified some cognitive (notably spatial abilities) and personality factors playing a major role in mental imagery-based BCI performances. They also revealed that both tactile feedback and social presence can improve BCI efficacy.

Finally, BCIs can be made more usable by being used for other applications than communication and control. To this end, we notably explored the use of BCIs for neuroergonomics, i.e., using brain signals to passively estimate some of the relevant user's mental states during human-computer interaction, in order to assess the ergonomic qualities of this interface. In particular, we showed that one can estimate mental workload during complex 3D manipulation and navigation tasks in order to assess or compare interaction techniques and devices. We have also been able to study stereoscopic displays by estimating visual comfort in EEG signals. Another usage of BCIs, that we found promising and useful, is real-time brain activity and mental state visualization. We designed a number of devices based on augmented reality and/or tangible interfaces to enable novice users to visualize their own brain activity or mental states in real-time, with potential applications in fields as wide and diverse as education, self-awareness or well-being.

Overall this work contributed novel methods and approaches to make EEG-based BCIs more usable, as well as new knowledge that could be used to further improve them in the future. This manuscript also proposes some perspectives and directions that could be worth exploring to that end. BCIs show a huge potential for research and applications, and we hope our research will contribute to turn this potential into realities.

Keywords Brain-Computer Interfaces (BCI), ElectroEncepahloGraphy (EEG), Classification, Signal Processing, Human Learning, Human-Computer Interaction, Teaching, Neuroergonomics

Host Laboratory Inria Bordeaux Sud-Ouest, 200 avenue de la vieille tour, 33405, Talence Cedex, France

Acknowledgements

I would like to start by thanking all the people who supported me and helped me in a direct or indirect way, in a small or in a big way, during these years of research (and of life) that are described in this manuscript. There are so many such great people that I was lucky enough to meet, that I know I will forget some of them, sorry in advance for that, and thank you!

I would like to thank all the members of my Habilitation committee, Benjamin Blankertz, Gernot Müller-Putz, Frédéric Dehais, Laurence Nigay, Maureen Clerc and Frédéric Alexandre. You are scientists that I respect a lot, who I truly find interesting, innovative and inspiring. I feel very honoured to have you all in my committee, so thank you for having taken the time to read and evaluate my work as well to come to my defence.

All the work presented in this manuscript is obviously not due to my sole contributions, but is the result of collaborations with many great people. Taken more or less from an historical order, thank you to Anatole Lécuyer, for having been a great PhD advisor and to continue to be a great collaborator and mentor still today. Thank you to all the colleagues from Rennes that I have been working with since my PhD, Yann Renard, Laurent George, Laurent Bonnet, Nicolas Brodu, Jonathan Mercier-Ganady, Maud Marchal, Bruno Arnaldi, Josef Legeny, Lorraine Perronnet, Christian Barrillot, Emilie Loup-Escande, Jussi Lindgren. It is always a pleasure, both intellectually and personally, to work with you. Brittany rocks!

Then, thank you to the colleagues from Singapore, notably Cuntai Guan and Kai Keng Ang, for the many things you taught me, and for having also opened my mind to new ways of doing research. Thank you as well to all the other colleagues for the BCI lab there who although we did not work directly together, still contributed interesting exchanges and discussions! So thank you notably to Haihong Zhang, Zheng Yang Chin, Chuanchu Wang and Kok Soon Phua.

Thank you then to Pascal Guitton and Martin Hachet, who welcomed me at Inria in Bordeaux, at that time in team Iparla. Thank you for having supported my application there, to have believed in me, and to have given me the freedom and support needed to do great research. Thank you to all the other members of the Iparla team at that time, notably Gaël Guennebaud, Romain Pacanowski, Xavier Granier, Pascal Barla, Christophe Schlick, Patrick

Reuter, Aurélie Cohé, Benoit Bossavit, Jérémy Laviole, Nicolas Mellado, Lucas Ammann, Nicolas Maréchal, Heqi Lu, Jiazhou Chen, Yan Li Liu, Hichem Barki, Marie Sanchez, and many others, for the ping-pong times, the parties, the scientific discussions and exchanges, and the incredible work atmosphere.

Thank you to all the amazing people I have been working with outside of Inria Bordeaux, in France and in the world. There are many many of them, so I am sure to forget some, but I'll try anyway, thank you to Laurent Bougrain, Jérémie Mattout, Théodore Papadopoulo, Sébastien Rimbert, Maureen Clerc, Florian Yger, Maxime Bérar, Julien Castet, Jean-Baptiste de la Rivière, Chi Tan Vi, Daniel Spelmezan, Stéphanie Fleck, Pierre Cutellic, Jan van Erp, Michaël Tangermann, Gerwin Schalk, Jonathan Brumberg, Robert Leeb, Christoph Guger, Moritz Grosse-Wentrup, Marco Congedo, Anton Andreev, Bernard N'Kaoua, Jean-Arthur Micoulaud-Franchi, Ricardo Chavarriaga, Sonja Kleih, Damien Coyle, Anton Nijholt, CS Nam, Paul McCullagh, Liam Maguire, Stephanie Lees, Amit Konar, Arijit Sinharay, Nicoletta Caramia, Masashi Sugiyama, Hideaki Toyama, Andreas Meinel, Josef Faller, Thorsten Zander, Laurens Krol, Lena Andreessen, Wojciech Samek, Sriram Subramanian. Thank you all for making my life so varied and exciting!

Then, some new thanks to Martin, for the Potioc adventure, that we created together, and for the many discussions, exchanges and new ideas that led, among many other things, to new and exciting research directions and results. Thank you to all the Potioc family, to the several Potioc generations that I already met, all of them being great. Potioc is really an incredible team to work in, with so many nice, smart and funny people. So thank you to Pauline Davignon, Thomas Hulin, Florian Larrue, Asier Marzo, Jacek Jankowski, Florent Berthaud, Luc Vercellin, Renaud Gervais, Anke Brock, Julia Chatain, Sol Roo, Fatma Ben Guefrache, Benoit Coulais, David Furio, Niko Katsakis, Maxime Duluc, Robin Gourdel, Nathalie Robin, Catherine Mégrat, Chrystel Plumejeau, Cécile Boutros, Boris Mansencal, Thibault Lainé, Pierre Antoine Cinquin, Damien Clergeaud, Marie Demengeat, Alexis Gay. You are such great colleagues and friends. I voluntarily left out the Potiocs I was closely working with on BCI, to give some extra special thanks. First thank to the many bachelor and master students I exploited, so thank you to Olivier Porpiglia, Loic Renaud, Leonard Pommereau, Aurélien Appriou, Joao Pedro Berti Ligabo, Julia Schumacher, Dennis Wobrock, Emilie Jahanpour, Suzy Teillet, Maxime Daniel, Benjamin Muzart and Manon Bonnet-Save, you did amazing jobs, especially given how little (or none) you were paid ;-). Thank you to Christian Mühl for your great work and friendship in your (too short) postdoc. Thank you Alison Cellard for your kindness and your many contributions to OpenViBE. Thank you to the new generation of PhD students Jelena Mladenovic and Lea Pillette for being so cool, for the great work you already did and the even greater work that I am sure you will do. Last but not least, and even most importantly a mega power thank you to my first two PhD students (now

PhD holders) Jeremy Frey and Camille Jeunet. I really feel extremely lucky to have had both of you as my first PhD students. You are both so perfect: kind, funny, easy to work with, smart, hardworking and crazy (this a positive thing in science :-)). Most of the work presented in this manuscript is actually your PhD works, so thousand of thank you! Some extra thanks to Camille, for your motivational help in writing this manuscript (your writing challenge was awesome), your help in obtaining my ERC (your slides were awesome) and for all the rest (you're awesome).

I am also lucky enough to be surrounded by great friends and family, who support and make my life so enjoyable everyday, which certainly positively contributed to the work presented here. Thank you to all my 'childhood' friends that I still see these days (childhood being defined here as until the end of high school :-)): 'Helmut', Sébastien, Olivier, Lucile, Clément, as well as to all the Rennes and CSOne tribe friends: Stéphanie, Stéfoo, Nicolas, Solenn, Peggy, Arnault, Fab, Marie, Grégoire, Patou, Bap, Laurent, Sarah, Gael, Aurélie, Pascale, Delphine, Olivier, Fred, Noémie, Gnafron, Anaëlle, Mathieu. Thank you as well to all the Singapore gang: Brahim, Dom, Lisa, Maryam, Guillaume, Mahnaz, Ahmed, Sepehr, Reza, Sidath, Atieh, Amir, Anousheh, Mahshad, Guoxian, Chris, Keng Peng, Andreas, Brice, Balint, Santhi, Omid, Nico, Agnés, Abdul, Nicholas. Big thanks as well to all the Bordeaux Swing Time and Koryo families (with a special thanks to Yann!). Thank you to the family, Maman, Papa, Muriel et Flora, for being such a loving and caring family, here as well I feel really lucky to have you, and although in an indirect way, you all contributed tremendously to the work in this manuscript. Finally, thank you very much to Morgane, my awesome girlfriend. Thank you for your support all these years, your help and your love. Thank you for having put up with me being away so often, for the late working hours and me disappearing in my thoughts to think - again - about my research. So thank you to still be by my side despite me being a shitty boyfriend that is not giving you the attention and time you deserve, I am really lucky to have you.

Altogether, I am so enjoying my work and my life, so thank you to all of you for making this an exciting adventure!

Contents

Introduction

A Brain-Computer Interfaces (BCI) can be defined as a system that translates the brain activity patterns of a user into messages or commands for an interactive application, this activity being measured and processed by the system [1, 2, 3]. A BCI user's brain activity is typically measured using ElectroEncephaloGraphy (EEG). For instance, a BCI can enable a user to move a cursor to the left or to the right of a computer screen by imagining left or right hand movements respectively [4, 5]. Since they make computer control possible without any physical activity, EEG-based BCIs have promised to revolutionize many applications areas, notably to enable severely motor-impaired users to control assistive technologies, e.g., text input systems or wheelchairs [6, 7, 8], as a rehabilitation device for stroke patients [9, 10], as new gaming input devices [11, 12] or to design adaptive human-computer interfaces that can react to the user's mental state [13], to name a few [14, 2].

Designing a BCI is a complex task which requires knowledge in multiple disciplines including computer science, engineering, signal processing, cognitive science, neuroscience and psychology. In order to use a BCI, two phases are generally required: 1) an offline training phase during which the system is calibrated and 2) the operational online phase in which the system can recognize brain activity patterns and translate them into commands for a computer. An online BCI system is a closed-loop, generally composed of seven main steps: brain activity pattern production, brain activity measurement, preprocessing, feature extraction, classification, decision and feedback:

1. **Brain activity pattern production** consists for the user in generating a specific, stable and distinct brain activity pattern so that it can be recognized in EEG signals and use as a control command. There are three main types of brain activity patterns that are used in EEG-based BCIs for communication and control: Event Related Desynchronization/Synchronization (ERD/ERS), Event Related Potentials (ERP) and Steady State Visual Evoked Potentials (SSVEP). ERD/ERS correspond to amplitude variations of EEG signals oscillations (a.k.a., oscillatory activity), i.e., amplitude changes in the power of EEG signals in certain frequency bands [15]. ERD/ERS can be observed during a number of different mental imagery tasks such as Motor Imagery (MI -

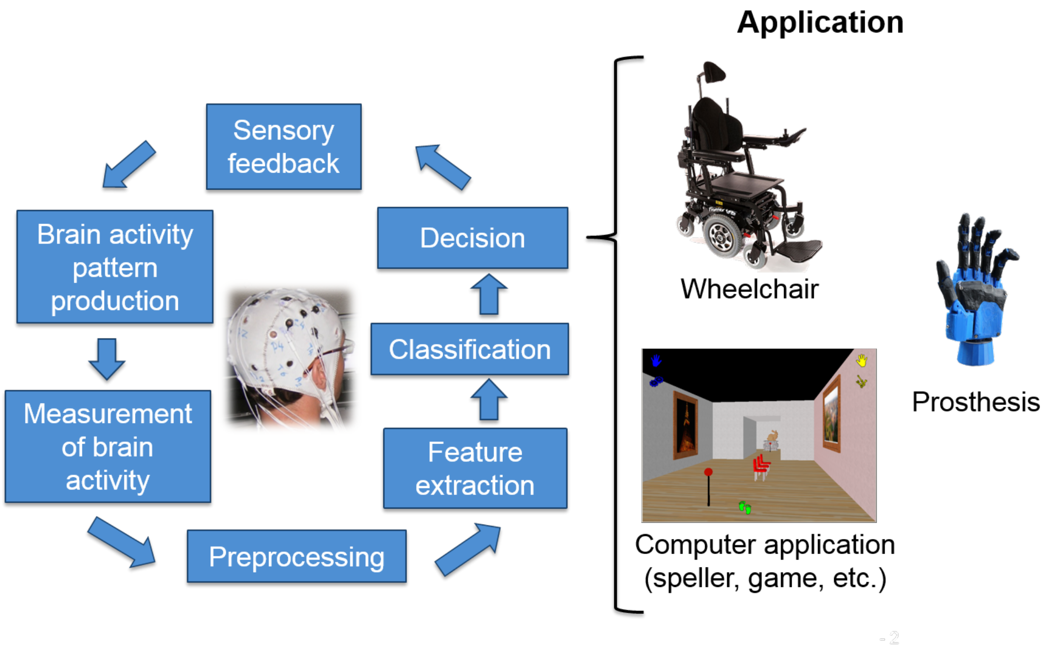


Figure 1: General architecture of an online brain-computer interface, with examples of applications.

limb movement imagination) [16] or mental rotation of a geometric figure [17, 18]. SSVEP are measured when the subject perceives and pays attention to a periodic stimulus such as a flickering picture. They are defined by an increase of the EEG power at the frequency of the stimulus and its harmonics [19]. An ERP is a brain response - characterized by specific temporal variations - due to some specific stimulus perceived by the BCI user. A typical ERP used for BCI design is the P300, which is a positive deflection of the EEG signal occurring about 300ms after the user perceived a rare and relevant stimulus [20].

2. **Brain activity measurement** allows to acquire the raw signals reflecting the user's brain activity [21]. Various types of sensors and measurements technologies can be employed, such as MagnetoEncephaloGraphy (MEG) [22], functional Magnetic Resonance Imaging (fMRI) [23] or intra-cortical electrodes (thus recording invasively) [24]. In this manuscript we focus on EEG as the measurement technique. Indeed, EEG is low-cost, portable, non-invasive and with a high temporal resolution [25]. As such, it is by far the most used recording modality so far, and the most likely to be used in practice outside laboratories [21].

3. **Preprocessing** consists in cleaning and denoising input data in order to enhance the relevant information contained in the raw signals [26]. This often consists in a set of spatial and/or spectral filters.

4. **Feature extraction** aims to describe the signals by a few relevant values called "features" [26]. These features can be, for instance, the power of the EEG over selected channels, and in specific frequency bands.
5. **Classification** assigns a class to a set of features extracted from the signals [27]. This class corresponds to the kind of brain activity pattern identified (e.g., left hand or right MI). Classification algorithms are known as "classifiers". Typical classifiers used for BCI include Linear Discriminant Analysis (LDA) or Support Vector Machines (SVM) [27].
6. **Decision** associates a command to the brain activity pattern (or sequence of patterns) identified in the user's brain signals, e.g., a recognized left hand movement could be translated into the command "move the cursor left".
7. **Feedback** is provided to the user to inform him/her about the recognized brain activity pattern. This aims to help the user modulate his/her brain activity and as such improve his/her control over the BCI [1]. Indeed, BCI is a skill that needs to be learned and refined [28].

This whole architecture is summarized in Figure 1. Currently, calibration is generally necessary in order to obtain a reliable BCI operation, and is generally done offline. In this stage, the classification algorithm is calibrated and the optimal features, and relevant sensors are selected. For this calibration, a training data set needs to be prerecorded from the user. Indeed, EEG signals are highly user-specific, and as such, most current BCI systems are calibrated specifically for each user. This training data set should contain EEG signals recorded while the user performed each mental task of interest several times, according to given instructions.

The BCIs described so far were systems for communication and control, in which the user was voluntarily sending mental commands to the application. These types of BCIs are known either as active BCI, when the user performs mental tasks (e.g., imagining movements), or as reactive BCI, when the users have to attend to stimuli (e.g., flickering visual images) [13]. There is yet another category of BCI: passive BCI, for which the mental state of the user is passively estimated, without any voluntary mental command from the user, to adapt the application in real-time to this mental state [13].

Since the design of the first real-time BCIs in the 90's [5, 29, 30, 31], the BCI field has grown tremendously, and now involves hundreds of laboratories and companies around the world [32], and has become a more mature field of research and technology [33, 2]. The BCI field have witnessed many recent innovations, including hybrid BCIs [34, 35], the exploration of new brain recording modalities such as functional Near Infrared Spectroscopy (fNIRS) [36, 37, 38] or ElectroCorticoGraphy (ECoG) [39], the availabilities on the market of consumer grade EEG and BCI systems such as OpenBCI (`openbci.com`), more robust EEG signal processing algorithms [40], open-source BCI softwares

[41], new BCI applications such as stroke rehabilitation [42], the development of passive BCI [13, 43], mobile BCI [44, 45] or using BCI as a new tool for research [46], among many others.

Despite this promising potential, the BCI revolution has not yet been delivered, and BCIs are still barely used outside research laboratories [2]. The main reason for this failed promise is the substantial lack of reliability and robustness of current BCIs [14, 2]. In particular, BCIs too often fail to correctly decode (i.e., recognize) the users' mental commands [47, 48, 49]. For example, the study in [47] showed that less than 80% of the users' mental commands were correctly decoded for more than 80% of the users (among 99 naive users), for a BCI based on only two MI tasks as mental commands. Similarly, in another study almost a decade later, with 80 users, the average rate of correct command decoding (a.k.a. classification accuracy) was still only 74.4%, also for two MI tasks [50]. Such performances are even lower with motor-impaired users [51, 52]. Moreover, it is estimated that 10% to 30% of BCI users, depending on the BCI type, cannot control the system at all (the so-called BCI illiteracy/deficiency) [53]. Moreover, while current EEG-based BCIs are reasonably stable in laboratory conditions, their performance decreases significantly when confronted to real-word, complex environments, over long periods of time, or when the user is moving [54, 55, 44]. They are very sensitive to noise, e.g., user motions or environmental magnetic or electric noises [55], as well as to the non-stationarity of EEG signals [56, 57]. Indeed, a BCI calibrated in a given context is very likely to have much lower performances when used in another context [58, 54]. EEG-based BCIs also generally require relatively long calibration time [59]. This is due to the need to collect numerous training EEG examples from each target user, to calibrate the BCI specifically for this user, to maximize performances [59]. Finally, many BCIs, particularly the active ones, also require long to very long human training [28, 60]. A number of challenges therefore still need to be tackled by the research community to yield robust, practical EEG-based BCIs.

To summarize, in Human-Computer Interaction (HCI) terms, BCIs suffer from a poor usability [61]. This means that they have both a low efficacy - they often send an erroneous command - and a low efficiency - they require a long time to be setup, calibrated and used. Thus, for EEG-based BCIs to fulfill their promises and leave laboratories to be used in practice, we - BCI researchers - need to dramatically increase their usability. Ideally, we need to make them more reliable, i.e., they should be able to recognize correctly the users' mental commands, whatever the user and the context, at all time. We also need to make BCI fast and easy to setup, to calibrate and to learn how to use. The contributions presented in this habilitation thesis aim at addressing these objectives. They comprise our research works since 2009 targeted at improving the usability of EEG-based BCIs. To do so, we explored three different research directions, 1) at the EEG signal processing and classification level, to robustly decode EEG signals and translate them into commands, 2) at the user training level, to ensure that users can learn to control a BCI efficiently and effectively, and 3) at the usage level, to explore novel applications of BCIs for which the current reliability can already be useful. For the first two points above, we focus on active, oscillatory activity-based BCIs using mental imagery tasks. Indeed, 1) they can be rather natural and intuitive to use, since their users "just" have to imagine a specific task to send control commands, 2) they do not require any stimulus (contrary to reactive BCIs), which enables the users to devote all their sensory attention to the feedback of the BCI and/or to their external environment, 3) they can be used in a self-paced way, that is, the user can initiate the task (and hence the command) at will, and 4) as we will see in Chapter 2, we believe there is much more room for improvement with such types of BCIs as compared to reactive BCIs (P300, SSVEP), by improving how users learn to control the system.

Our first set of contributions is presented in Chapter 1, which deals with EEG signal processing tools. In this chapter we proposed a number of methods to improve BCI reliability, despite EEG signal variability, poor signal-to-noise ratio and high sensitivity to artifacts, as well as to reduce BCIs calibration times. More precisely, we complemented traditionally used features by exploring alternative representations of EEG signals. We also explored and designed regularized spatial filters to learn more robust and stable features. Finally, we propose algorithms to reduce BCIs calibration times, i.e., to calibrate BCIs with as few examples of EEG signals from the target user as possible, by re-using data from previous users or by generating artificial EEG signals. Altogether, these methods enabled an increased BCI classification accuracy, i.e., efficacy, and BCI calibration with much less data than standard approaches do, thus improving their efficiency.

In Chapter 2, rather than improving EEG signal processing alone, we advocate that BCIs can also be made more usable by guiding users to efficiently learn BCI control mastery. Indeed, BCI control is known to be a skill that needs to be learnt [28]. A study of models and guidelines from educational sciences enabled us to identify many theoretical limitations of current standard BCI training approaches, thus highlighting the need for alternative ones. In particular, educational sciences recommend to train people with adapted and adaptive training tasks, using explanatory feedback in motivating environments. In contrast standard BCI training protocols are commonly fixed, repetitive, rather boring and provide purely corrective feedback. To address these limitations, we studied what kind of users manage to use a BCI and why. We also explored new feedback types, in particular richer feedback, multi-user feedback and tactile feedback to help users to learn BCI control skills more efficiently. Overall, our studies identified some cognitive (notably spatial abilities) and personality factors playing a major role in mental imagery-based BCI performances. They also revealed that both tactile feedback and social presence can improve BCI efficacy.

Finally, in Chapter 3, we show how BCIs can be made more usable by being used for other applications than communication and control. To this end, we notably explored the use of BCIs for neuroergonomics, i.e., using brain signals to passively estimate some of the relevant user's mental states during human-computer interaction, in order to assess the ergonomic qualities of this interface. In particular, we showed that one can estimate mental workload during complex 3D manipulation and navigation tasks in order to assess or compare interaction techniques and devices. We have also been able to study stereoscopic displays by estimating visual comfort in EEG signals. Another usage of BCIs, that we found promising and useful, is real-time brain activity and mental state visualization. We designed a number of devices based on augmented reality and/or tangible interfaces to enable novice users to visualize their own brain activity or mental states in real-time, with potential applications in fields as wide and diverse as education, self-awareness or well-being.

These contributions and the outline of this manuscript are summarized on Figure 2. Overall these works contribute novel methods and approaches to make EEG-based BCIs more usable, as well as new knowledge that could be used to further improve them in the future.

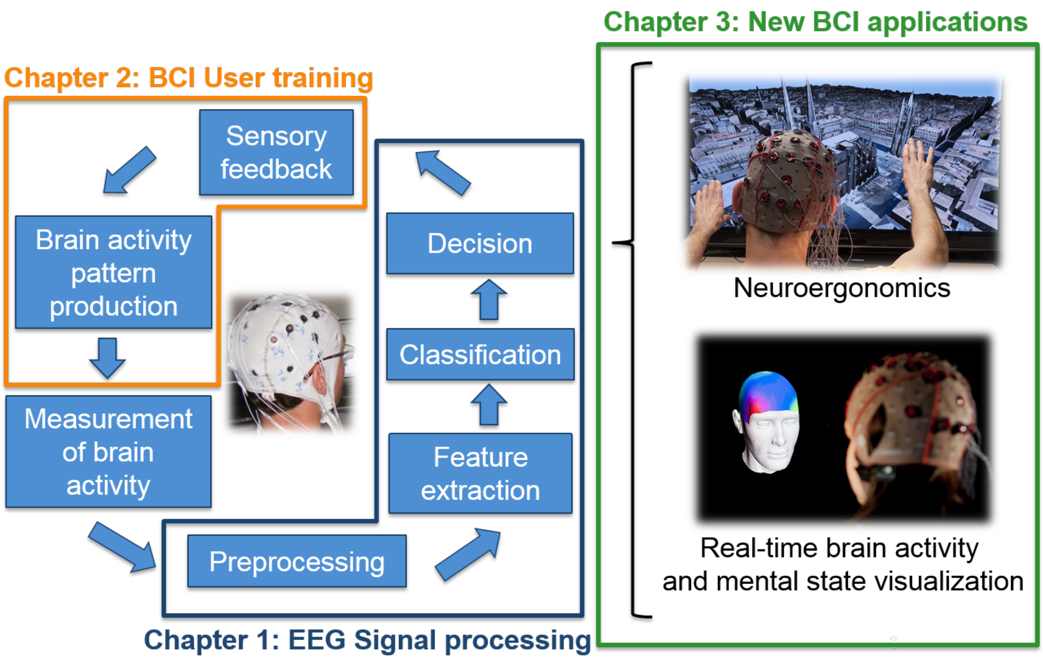
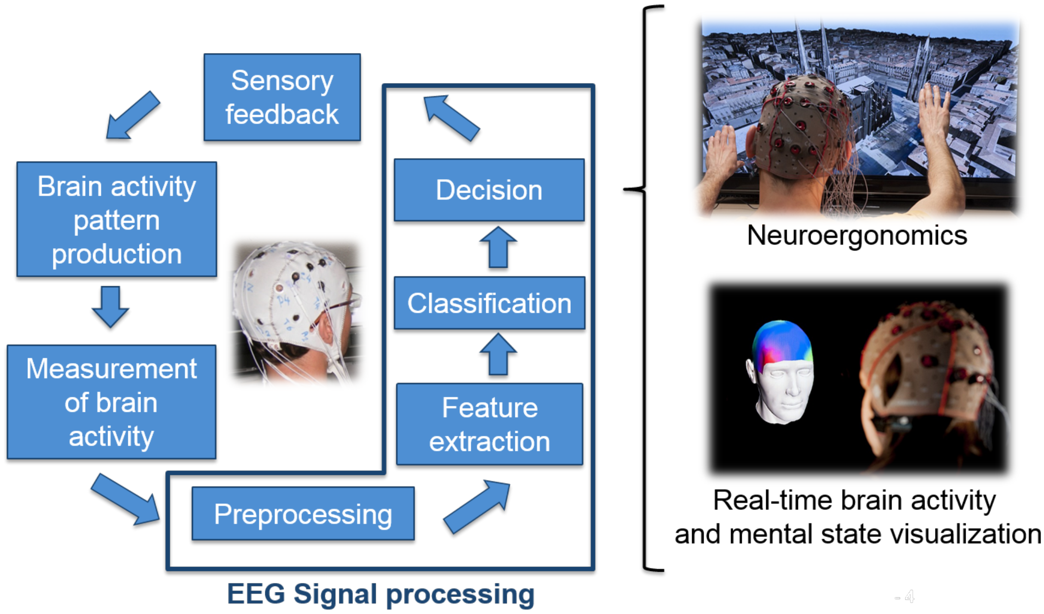


Figure 2: The contributions of this habilitation thesis and their corresponding chapters.

Chapter 1

EEG signal processing tools for robust BCI design with minimal calibration time



Selected related Publications:

- F. Yger, M. Berar, F. Lotte, *Riemannian approaches in Brain-Computer Interfaces: a review*, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2016 - accepted with minor revisions
- F. Lotte, *Signal processing approaches to minimize or suppress calibration time in oscillatory activity-based Brain-Computer Interfaces*, Proceedings of the IEEE, vol. 103, no. 6, pp. 871-890, 2015
- F. Yger, F. Lotte, M. Sugiyama, *Averaging covariance matrices for EEG*

signal classification based on the CSP: an empirical study, Proc. EUSIPCO, pp. 2721 - 2725, 2015
- N. Caramia, F. Lotte, S. Ramat, *Optimizing spatial filter pairs for EEG classification based on phase synchronization*, Proc. International Conference on Audio, Speech and Signal Processing (ICASSP'2014), pp. 2049-2053, 2014
- F. Lotte, *A new feature and associated optimal spatial filter for EEG signal classification: Waveform Length*, International Conference on Pattern Recognition (ICPR'2012), pp. 1302-1305, 2012
- N. Brodu, F. Lotte, A. Lécuyer, *Exploring Two Novel Features for EEG-based Brain-Computer Interfaces: Multifractal Cumulants and Predictive Complexity*, Neurocomputing, vol. 79, no. 1, pp. 87-94, 2012
- F. Lotte, C.T. Guan, *Regularizing Common Spatial Patterns to Improve BCI Designs: Unified Theory and New Algorithms*, IEEE Transactions on Biomedical Engineering, vol. 58, no. 2, pp. 355-362, 2011
- F. Lotte, C.T. Guan, *Learning from other Subjects Helps Reducing Brain-Computer Interface Calibration Time*, International Conference on Audio, Speech and Signal Processing (ICASSP'2010), pp. 614-61 7, 2010
- F. Lotte, C.T. Guan, K.K. Ang, *Comparison of Designs Towards a Subject-Independent Brain-Computer Interface based on Motor Imagery*, 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 4543-4546, 2009

Scientists that I (co-)supervised for this work:

- Visitng PhD student:
  - Nicoletta Caramia
- Engineer:
  - Alison Cellard
- Master student:
  - Joao-Pedro Berti Ligabo

1.1 Introduction

As mentioned previously, the poor usability of BCI is due in part to the poor accuracy with which mental commands are recognized, as well as to the long calibration times that are necessary to calibrate BCI signal processing algorithms. In this chapter we propose new EEG signal processing algorithms to address these limitations. In particular, we first propose contributions to increase the mental command recognition accuracy of BCIs by exploring additional features to complement traditional ones, as well as by proposing robust spatial filters that accommodate noisy signals and non-stationarity. Then, to reduce BCI calibration times, we propose different algorithms exploiting data from other subjects, statistical regularization and artificial data generation. It should be reminded that all these algorithms are designed for BCIs based on oscillatory activity such as MI-based BCIs.

This chapter is organized as follows: Section 1.2 first presents the standard design that is currently used to design BCIs based on oscillatory activity. Then, the following sections present our new signal processing algorithms that are improvement on this standard design. In particular Section 1.3 presents the alternative features we explored and designed. Then Section 1.4 presents robust variants of the standard spatial filtering algorithm for oscillatory BCIs. Finally, Section 1.5 presents different algorithms to reduce or suppress calibration times with such designs.

1.2 Standard oscillatory activity-based BCI design

A typical oscillatory activity-based BCI is designed around two main algorithms: the Commom Spatial Patterns (CSP) algorithm to optimize spatial filters [1] and the Linear Discriminant Analysis (LDA) algorithm for classification. The CSP algorithm aims at learning spatial filters such that the variance of the spatially filtered signals is maximized for one class (e.g., one mental imagery task) and minimized for the other class. Since the variance of a band-pass filtered signal corresponds to the band-power of the signal in that band, CSP optimizes spatial filters that lead to optimally discriminant band-power features [62]. This is particularly interesting and relevant for the design of oscillatory activity-based BCI since such BCI exploit changes in EEG oscillations amplitude, i.e., changes in the EEG signals band power. More formally, optimizing CSP spatial filters w (w being a weight vector [2]) consists in extremizing

1. A spatial filter is a (usually linear) combination of the original channels. Performing spatial filtering helps to overcome the EEG spatial blurring that occurs due to EEG signals traveling through the skull and scalp [62]

2. In this manuscript, all vectors are assumed to be row vectors

the following function:

$$J_{CSP}(w) = \frac{wC_1w^T}{wC_2w^T} \tag{1.1}$$

where C_i is the average spatial covariance matrix of the band-pass filtered EEG signals from class i, and T denotes transpose. Typically, these spatial covariance matrices are obtained by computing the spatial covariance matrix C_i^j from each trial T_i^j from class i, and then averaging them:

$$C_i = \frac{1}{N_i}\sum_{j}^{N_i} C_i^j = \frac{1}{N_i}\sum_{j}^{N_i} T_i^j(T_i^j)^T \tag{1.2}$$

with N_i the number of trials in class i and $T_i^j \in \mathbb{R}^{C\times S}$ is the j^{th} EEG trial from class i, with S the number of samples in a trial, and C the number of channels. Note that the EEG signals are assumed here to be band-pass filtered and thus to have zero mean. This optimization problem is solved by Generalized Eigen Value Decomposition (GEVD) of the two matrices C_1 and C_2 [62]. The spatial filters which maximize/minimize $J_{CSP}(w)$ are the eigen vectors corresponding to the largest and smallest eigen values of this GEVD, respectively. Once the filters w are obtained, CSP feature extraction consists in filtering the EEG signals using the w and then computing the resulting signals variance. In other words, a feature f is computed as $f = log(wC_{ct}w^T)$, where C_{ct} is the current trial covariance matrix.

The LDA classifier uses a linear hyperplane to separate feature vectors from two classes [27]. The intercept b and normal vector a of this hyperplane are computed as follow:

$$a^T = C^{-1}(\mu_1 - \mu_2)^T \tag{1.3}$$

$$b = -\frac{1}{2}(\mu_1 + \mu_2)a^T \tag{1.4}$$

with μ_1 and μ_2 being the mean feature vectors for each class and C the covariance matrix of both classes. With LDA, for an input feature vector x, the classification output is $ax^T + b$. If this output is positive, the feature vector is assigned to the first class, otherwise it is assigned to the second class. The whole process is summarized in Figure 1.1.

1.3 Exploring alternative features

The standard BCI design presented above only makes use of a single feature type, namely Band Power (BP) features. There are, however, many other ways in which EEG signals can be described and analyzed, in order to represent various neurophysiological phenomena. In other words, since the current BCI performances are still far from being satisfactory, there is a need to explore and

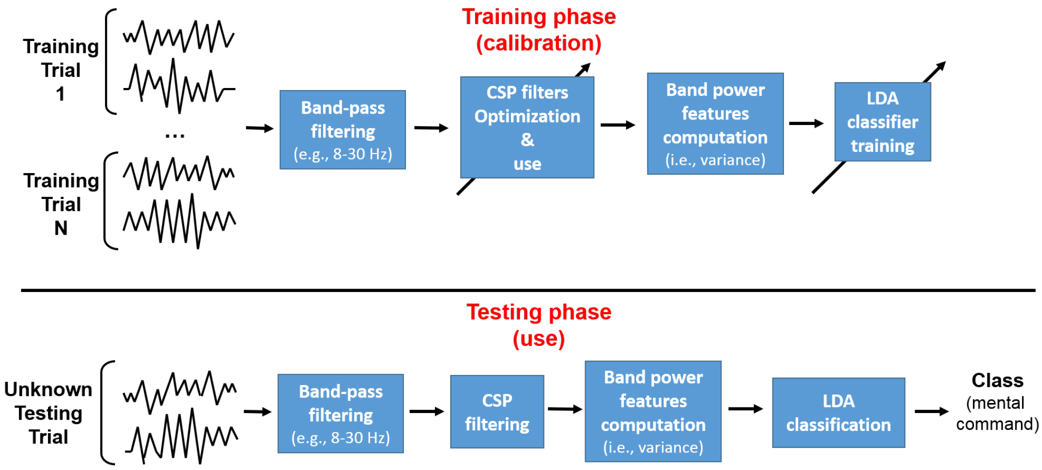


Figure 1.1: The training (a.k.a. calibration) and testing steps involved in the standard design of an oscillatory activity-based BCI. During calibration, both the CSP spatial filters and the LDA classifier are optimized based on training data. During testing (i.e., actual use of the BCI with new EEG signals), the CSP filters are used to spatially filter the signals from which band power features are computed. These features are classified using the optimized LDA to identify the mental task performed by the user.

to design alternative features, which the BCI community has stressed several times [63, 64]. In this section, we therefore present several alternatives features that we have explored and designed in order to improve BCI performances and robustness. In particular, we explored here 1) multifractal cumulants, which describes relationships between different frequency bands of a signal [65], 2) two measures of complexity of a signal, namely, its predictive complexity [65] and its waveform length [66], and 3) phase locking value, which is a measure of synchronization between different brain signals [67]. For the last two features, we also proposed spatial filters that aim at making such features as discriminant as possible.

1.3.1 Alternative features explored

Multifractal cumulants

MultiFractal Cumulants (MFC) can be seen as a statistic on inter-frequency band relations. This is particularly relevant for BCI as this information is generally ignored in current MI-based BCI designs. The multifractal approach adds in information about how the multiple bands relate together at each instant. The multifractal formalism is described in details in [68, 69]. The method we chose for extracting the multifractal spectrum is a discrete wavelet transform of the signal, out of which we extract the wavelet leader coefficients

[70]. Following the directions of [68] we then use the cumulants of the leaders as the features for classification. Interested readers can refer to [65] for details. In short, one view on multifractal analysis [71] is to relate the statistical properties of the signal $x(t)$ to analyse, and of a scaled version of it $x(at)$. In terms of frequency analysis, that scaling in time corresponds to a frequency shift.

Predictive complexity

The Predictive Complexity (PC) of a temporal signal describes the statistical complexity and predictive properties of the time series [72]. The information (quantified in bits) that is extracted this way measures how difficult it is to make an optimal prediction based on past information. It is null both for totally ordered and totally random systems, and increases in between. The intuitive idea behind this feature is to quantify the amount of information that is necessary to retain from the past of the series in order to be able to predict optimally the future of the series [73]. Doing so can be achieved using the concept of decisional states.

Informally, the idea behind the decisional states is to construct a Markovian automaton [74, 73] whose states correspond to taking the same decisions [72], according to a user-defined utility function. These decisions are those that one can take based on predictions of the future and their expected utility. The complexity of the series is then computed as the mutual information between the internal states of the Markovian automaton, and the series itself. The complexity is null for a very regular series, for example a constant series or a series where we always take the same decision: there is only a single state in the automaton. Similarly the complexity of a completely random series is also null: it can be modelled by successive independent draws from a fixed probability distribution, whose expected utility we take to make our decision. This leads again to a single Markovian automaton state, hence a null complexity. The complexity measure increases only for more complicated series with many internal states (i.e. many distinct probability distributions of what happens next, depending on what previously happened, leading to different decisions). For the sake of conciseness, the formal description is left out of this manuscript, but the interested reader can refer to [65] for the details.

Waveform length and associated spatial filter

The Waveform Length (WL) measures the length of a given waveform, which is also a measure of the signal complexity [75]. WL was initially designed to classify ElectroMyoGraphy (EMG) signals, and has been shown to be one of the most robust features for this task [75]. As both EEG and EMG measure an electrical signal resulting from the activity of neuron populations (cortical neurons for EEG, motor ones for EMG), it seemed promising to explore whether a feature that can successfully classify EMG could also successfully classify EEG. Formally, The WL of an EEG signal x is:

$$wl = \log(\sum_{i=1}^{N-1} |x_{i+1} - x_i|) = \log(\sum_{i=1}^{N-1} |\Delta x_i|) = \log(\|\Delta x\|_1)$$

with $|x|$ being the absolute value of x, and $\|\|_1$ the $l1$-norm. This feature measures the cumulative length of the EEG signal analyzed. To maximize the efficiency of this feature for EEG classification, it should be extracted after appropriate spatial filtering, in the same way as band power efficiency is maximized by CSP spatial filtering. Therefore, we also propose a spatial filter that is optimal for classification based on WL features. We denote this new spatial filter WL Optimal Spatial Filter (WOSF). In order to derive such an algorithm, we have to find spatial filters w which maximize the waveform length of spatially filtered EEG signals from one class, while minimizing it for the other class. Formally, this means extremizing the following function:

$$J_{WOSF1}(w) = \frac{\|w^T X_1^{2:N} - w^T X_1^{1:(N-1)}\|_1}{\|w^T X_2^{2:N} - w^T X_2^{1:(N-1)}\|_1} = \frac{\|w^T \Delta X_1\|_1}{\|w^T \Delta X_2\|_1}$$

with $\Delta X = X^{2:N} - X^{1:(N-1)}$ and $X^{i:j}$ being the signal matrix X with only rows i to j, i.e., with only EEG samples from indexes i to j. Unfortunately, the $l1$-norm is not differentiable. This makes the optimization of J_{WOSF1} inconvenient, iterative, complex and computationally expensive. Therefore, we decided to optimize the spatial filters using the $l2$-norm rather than the $l1$-norm, which, as we will see later on, leads to a closed-form and computationally efficient solution, similar to that of CSP. Thus, our objective function becomes:

$$J_{WOSF2}(w) = \frac{\|w^T \Delta X_1\|_2}{\|w^T \Delta X_2\|_2} = \frac{w^T \Delta X_1^T \Delta X_1 w}{w^T \Delta X_2^T \Delta X_2 w} = \frac{w^T D_1 w}{w^T D_2 w}$$

with $D_i = \Delta X_i^T \Delta X_i$. This is, as for CSP, a generalized Rayleigh quotient, and as such, the spatial filters which maximizes or minimizes J_{WOSF2} are the eigenvectors corresponding to the largest and lowest eigenvalues obtained by GEVD of matrices D_1 and D_2. As for CSP, the ΔX_i matrices used in practice are the average ΔX matrices computed for each trial of class i. Once the WOSF spatial filters are obtained, extracting feature wl_i for the i^{th} spatial filter w_i is simply achieved as $wl_i = \log(\|w_i^T \Delta X\|_1)$. For further details the interested reader can refer to [66].

Phase locking value and associated spatial filter

The phase locking value (PLV) measures the synchronization between the signals from two different EEG signals [76, 77]. Such features have indeed been shown to be promising to classify EEG signals for BCI [78, 79]. It can be computed as follows:

$$S_{PLV} = \frac{1}{N} \sum_{(k=1)}^{N} e^{j|\phi^1(k) - \phi^2(k)|} \tag{1.5}$$

Where N is the number of samples in the considered time window, ϕ^1 and ϕ^2 are the phase values of the two signals that we want to compare [77]. The computed S_{PLV} is a number between 0 and 1 that reflects how the two signals are synchronized to each other.

This feature is not new for BCI in itself. However, as for WL, such feature could benefit from the use of dedicated spatial filters, to reduce the volume conduction effect, and to find two brain areas (each obtained using a dedicated spatial filter) whose signals synchronization is maximally different between classes. We therefore proposed an algorithm to optimize these two spatial filters w_1 and w_2 in order to maximize the resulting PLV feature discriminative power. To do so, the idea was to maximize the difference in S_{PLV} of the two spatially filtered signals between the two classes, which can be done by maximizing the following function:

$$Diff = |S_{PLV}(w_1^t X_1, w_2^t X_1) - S_{PLV}(w_1^t X_2, w_2^t X_2)| \tag{1.6}$$

where X_i is EEG traces recorded for class i (X_i in $\mathbb{R}^{(Nc \times Ns \times Nt)}$ with Nc the number of channels, Ns the number of samples and Nt the number of trials per class). In other words, this functional amounts to optimizing spatial filters such that the resulting S_{PLV} value is maximally different between the two classes. To avoid obtaining two identical spatial filters, which would be useless (a signal is necessarily synchronized with itself), we proposed to enforce the orthogonality of the two spatial filters by introducing a regularization term expressed as:

$$Ort = \frac{|w_1^t - w_2 * Diff|}{dim(w_2)} \tag{1.7}$$

In this way the orthogonality value was scaled according to the value computed in Diff. Combining the two members gives the final functional which should be maximized: $fitness = Diff - \alpha \times Ort$ where α was a parameter between 0 and 1 indicating how much the second member weighs, i.e., how much we want to enforce the spatial filter orthogonality. This objective function can be optimized with various algorithms. In our experiments, we used genetic algorithms to do so (see [67] for details), but these are by no means the only possible solution, and other optimization algorithms may even prove faster and more stable to do so. In the following we denote as PLV-SP the features obtained by computing PLV features on the spatially filtered signals.

1.3.2 Evaluation

These different alternative features were evaluated on different EEG data sets, as reported in [65, 66, 67]. We briefly describe them below. The details about how all these data sets were preprocessed for each feature can be found in [65, 66, 67].

Data set III, BCI competition II (DSIIIBCICompII)

Data set III from BCI competition II [80] contains 280 trials of left and right hand Motor Imagery (140 trials per class) from 1 subject. EEG were recorded using the C3, C4 and Cz electrodes, however, for the purpose of this evaluation, we used only the C3 and C4 electrodes as recommended in [81]. More details about this data set can be found in [80].

Data set IIIb, BCI competition III (DSIIIbBCICompIII)

Data set IIIb from BCI competition III originally consists of EEG signals from 3 subjects. However, for the purpose of this study, only subjects labeled S4 and X11 were used. Indeed, EEG signals for subject O3VR were recorded using a different protocol and the data file provided online contained erroneously duplicate signals[3]. For both subjects, both their training and their testing sets were composed of 540 trials of left versus right motor imagery. The data was recorded using electrodes C3 and C4. See [82] for details.

Data Set IIa, BCI competition IV (DSIIaBCICompIV)

Data set IIa from BCI competition IV [83, 84], comprises EEG signals from 9 subjects who performed left hand, right hand, foot and tongue MI. The EEG signals were recorded using 22 EEG channels over the sensorimotor cortex. For the purpose of these studies, only EEG signals corresponding to left and right hand MI were used. A training and a testing set were available for each subject, both containing 72 trials for each class.

Data set IIb, BCI competition IV (DSIIbBCICompIV)

Data set IIb, BCI competition IV [83, 85], comprises data from 9 subjects who performed left hand and right hand MI. EEG signals were recorded using 3 bipolar channels around C3, Cz and C. The data contains ocular artifacts that interfere with the brain signals. There were 200 trials per class for training, and 120 trials per class for testing.

3. See http://www.bbci.de/competition/iii/desc_IIIb_ps.html for details

OpenViBE data set (DSOV)

This data set comprises EEG signals from a single subject for which 560 trials of motor imagery (280 trials per class, left vs right imagined hand movement) were recorded over a 2 week period. Half of the trials (randomly selected from all experiments over time) is used for training, the remaining half for testing. EEG data was recorded using the following electrodes: C3, C4, FC3, FC4, C5, C1, C2, C6, CP3, CP4, with a nose reference electrode. See [65] for details.

Homemade Mental Rotation data set (DSMR)

This data set was collected in-house from 6 subjects during Mental Rotation (MR) tasks. The protocol used was similar to that of data set DSIIaBCICompIV, except that on cue presentation, instead of performing MI tasks, subjects were instructed to either imagine continuous rotations of a 3D geometric figure displayed on screen or to relax while fixating a dot displayed in the screen center. EEG were collected using 15 electrodes (C3, C1, Cz, C2, C4, F3, F1, Fz, F2, F4, P3, P1, Pz, P2, P4). Each subject participated to 4 runs, a run comprising 20 trials from each class (relax and MR), except subject B3, who participated to 3 runs only, due to fatigue. See [66] for details.

Methods

These different features were assessed on the data sets above, using classical classifiers, either a LDA for the MFC, PC and WL features [65, 66], or a Support Vector Machine (SVM) for the PLV features [67]. If there were hyperparameters to set up for these features, they were either selected using cross-validation on the training set, or a fixed default value was used, the same for all subjects (see [65, 66, 67] for details). The features were evaluated individually, as well as together with their corresponding optimal spatial filters, if any (i.e., with WOSF and PLV-SP for WL and PLV respectively). They were also compared with the classical Band Power (BP) features, as well as with CSP and BP features in case they were also using spatial filters (for WOSF and PLV-SP). Finally, they were also combined with the classical BP features, with or without CSP spatial filtering, to assess whether they could bring additional information, not captured by the standard design, i.e., BP+CSP. For this combination the different features were concatenated together into a larger feature vector before being used as input to the classifier. This was used for the WL and PLV features. For the MFC and PC features, the dimensionality was higher, since no spatial filters were used to reduce the dimensionality. Thus, in order to reduce the negative influence from the curse-of-dimensionality, for these features, their combination was achieved by first training an LDA classifier for each feature type, and then using a weighted average of these different classifier output to make the final decision. The weight used for each classifier was the Fisher ratio of this classifier on the training set [65].

1.3.3 Results

The classification performances obtained by the different features, on different data sets, alone, or in combination with the classical BP features, are reported below, in Table 1.1 for the MFC and PC features, in Table 1.2 for the WL features and in Table 1.3 features.

Table 1.1: Average classification accuracy (%) obtained with the MultiFractal Cumulant (MFC), Predictive Complexity (PC) and Band Power(BP) features, alone or in combination

Paper	Data sets	Number of subjects	BP	MFC	PC	BP +MFC +PC
[65]	DSIIIBCICompII DSIIIbBCICompIII DSIIbBCICompIV DSOV	13	78.9	75.8	71.4	**80.3**

Table 1.2: Average classification accuracy (%) obtained with the Waveform Length (WL) feature, with or without its Optimal Spatial Filter (WOSF), as compared to or combined with Band Power (BP), with or without Common Spatial Patterns (CSP) spatial filtering

Paper	Data sets	Number of subjects	BP	WL	CSP	WOSF	CSP + WOSF
[66]	DSIIaBCICompIV DSMR	15	68	66.7	77	78.7	**80.1**

Table 1.3: Average classification accuracy (%) obtained with the Phase Locking Value (PLV) features, with or without its Optimal Spatial Filter (PLV-SP), as compared to or combined with Band Power features and Common Spatial Patterns (CSP) spatial filtering

Paper	Data sets	Number of subjects	PLV	CSP	PLV-SP	PLV-SP + CSP
[67]	DSIIaBCICompIV	9	64.1	75.8	73.9	**78**

1.3.4 Conclusion on Alternative features

Several interesting insights can be obtained from the results. First, it can be observed that all alternative features lead to classification accuracies that are clearly better than chance, which shows such features do carry discriminatory information about the different mental imagery tasks. Second, it shows that appropriate spatial filtering does substantially improve the classification performances, for each feature (BP, WL and PLV). This confirms the importance of spatial filtering for BCI design, and that our newly designed spatial filters for the alternative features are relevant and useful. Finally, results show that the best feature type still is BP. However, the best overall performances are always obtained when combining the new feature types with the BP features (with or without spatial filtering). This superiority is statistically significant ($p < 0.05$). This means that the alternative features we have explored actually extract a different information than the BP features, which can thus complement such features and lead to improved BCI performances.

1.4 Robust spatial filter design

As presented in Section 1.2, spatial filtering, and in particular the CSP algorithm, is a key element in oscillatory activity-based BCI design. Indeed, the CSP algorithm has numerous advantages: first, it leads to high classification performances. CSP is also versatile, since it works for any ERD/ERS BCI. Finally, it is computationally efficient and simple to implement. Altogether this makes CSP one of the most popular and efficient approach for BCI based on oscillatory activity [62]. Nevertheless, despite all these advantages, CSP is not exempt from limitations. In particular, CSP has been shown to be non-robust to noise, to non-stationarities and prone to overfitting (i.e., it may not generalize well to new data) when little training data is available [86, 87, 88]. Therefore, MI-based BCI can be made more usable, and in particular more effective, by making the CSP algorithm more robust to noise, non-stationarities and small sample settings. In this section, we propose two approaches to do so. The first one consists in regularizing the CSP algorithm, i.e., in using prior knowledge to guide the spatial filter optimization process. The second approach consists in robustly averaging covariance matrices, on which CSP is based, to accommodate noise and outliers.

1.4.1 Regularizing CSP

One way to make CSP robust and stable with limited and noisy training data is to integrate prior knowledge into its optimization algorithm. Such knowledge could represent any information we have about what should be a good spatial filter for instance. This can be neurophysiological prior, data

(EEG signals) or meta-data (e.g., good channels) from other subjects, etc. This knowledge is used to guide and constraint the CSP optimization algorithm towards good solutions even with noise, limited data and non-stationarities [89]. Formally, this knowledge is represented in a regularization framework that penalizes unlikely solutions (i.e., spatial filters) that do not satisfy this knowledge, therefore enforcing it. Similarly, prior knowledge can be used to stabilize statistical estimates (here, covariance matrices) used to optimize the CSP algorithm. Indeed, estimating covariance matrices from few training data usually leads to poor estimates [90]. We proposed a framework to design such a Regularized CSP (RCSP).

A regularization framework for CSP

Formally, a Regularized CSP (RCSP) can be obtained by maximizing both equation 1.8 and 1.9:

$$J_{RCSP1}(w) = \frac{w\tilde{C}_1 w^T}{w\tilde{C}_2 w^T + \lambda P(w)} \tag{1.8}$$

$$J_{RCSP2}(w) = \frac{w\tilde{C}_2 w^T}{w\tilde{C}_1 w^T + \lambda P(w)} \tag{1.9}$$

$$\text{with } \tilde{C}_i = (1-\gamma)C_i + \gamma G_i \tag{1.10}$$

In these equations, $P(w)$ is the penalty term that encodes the prior knowledge. This is a positive function of the spatial filter w, whose value will increase if w does not satisfy the knowledge encoded. Since the filters are obtained by maximizing J_{RCSPi}, this means that the numerator (which is positive) must be maximized and the denominator (which is also positive) must be minimized. Since $P(w)$ is positive and part of the denominator, this means that $P(w)$ will be minimized as well, hence enforcing that the spatial filters w satisfy the prior knowledge. Matrix G_i is another way of using prior knowledge, in order to stabilize the estimates of the covariance matrices C_i. If we have any idea about how these covariance matrices should be, this can be encoded in G_i in order to define a new covariance matrix $\tilde{C}_i$ which is a mix of the matrix C_i estimated on the data and of the prior knowledge G_i. We will present below what kind of knowledge can be encoded in $P(w)$ and G_i. λ and γ are regularization parameters controlling the amount of the regularization.

For the penalty term $P(w)$, a kind of knowledge that can be used is spatial knowledge. For instance, from a neurophysiological point of view, we know that neighboring neurons tend to have similar functions, which supports the idea that neighboring electrodes should measure similar brain signals (if the electrodes are close enough to each other), notably because of the smearing effect. Thus neighboring electrodes should have similar contributions in the spatial filters. In other words, spatial filters should be spatially smooth. This can be enforced by using the following penalty term:

$$P(w) = \sum_{i,j} Prox(i,j)(w_i - w_j)^2 \tag{1.11}$$

Where $Prox(i,j)$ measures the proximity of electrodes i and j, and $(w_i - w_j)^2$ is the weight difference between electrodes i and j, in the spatial filter. Thus, if two electrodes are close to each other and have very different weights, the penalty term $P(w)$ will be high, which would prevent such solutions to be selected during the optimization of the CSP. We denote this CSP variant as a Spatially Regularized CSP (SRCSP). Another knowledge that can be used is that for a given mental task, not all the brain regions are involved and useful. As such, some electrodes are unlikely to be useful to classify some specific mental tasks. This can be encoded in $P(w)$ as well:

$$P(w) = wDw^T \quad with \quad D(i,j) = \begin{cases} channel \quad i \quad \text{“}uselessness\text{''} \quad if \quad i = j \\ 0 \qquad\qquad\qquad\qquad \text{otherwise} \end{cases} \tag{1.12}$$

Basically, the value of $D(i,i)$ is the penalty for the i^{th} channel. The higher this penalty, the less likely this channel will have a high contribution in the CSP filters. The value of this penalty can be defined according to neurophysiological prior knowledge for instance, large penalties being given to channels unlikely to be useful and small or no penalty being given to channels that are likely to genuinely contribute to the filter. However, it may be difficult to precisely define the extent of the penalty from the literature. Another alternative is the use data previously recorded from other subjects. Indeed, the optimized CSP filters already obtained from previous subject give information about which channels have large contributions on average. The inverse of the average absolute weight of each channel in other subjects CSP filters can be used as the penalty, hence penalizing channels with small average contribution [89]. We denote this CSP variant as a Weighted Tikhonov Regularized CSP (WTRCSP). Note that other regularization terms can be used [89]. We will also see in Section 1.5 how regularization of the covariance matrices estimate can be used to reduce calibration time.

Evaluation

We evaluate these regularized CSP approaches on 3 publicly available data sets from BCI competitions: Data set IIa from BCI competition IV (DSIIaBCICompIV - already described in Section 1.3.2), Data set IIIa from BCI competition III and Data set IVa from BCI competition II. These two data sets not described before are described here after.

Data set IVa, BCI competition III (DSIVaBCICompIII) Data set IVa [91], from BCI competition III [82], contains EEG signals from 5 subjects, who performed right hand and foot MI. EEG were recorded using 118 electrodes. 280 trials were available for each subject, among which 168, 224, 84, 56 and 28 composed the training set for subject A1, A2, A3, A4 and A5 respectively, the remaining trials composing their test set.

Data set IIIa, BCI Competition III (DSIIIaBCICompIII) Data set IIIa [92], from BCI competition III [82], comprises EEG signals from 3 subjects who performed left hand, right hand, foot and tongue MI. EEG signals were recorded using 60 electrodes. For the purpose of this study, only EEG signals corresponding to left and right hand MI were used. A training and a testing set were available for each subject. Both sets contain 45 trials per class for subject B1, and 30 trials per class for subjects B2 and B3.

To classify EEG signal features (i.e., band power features), after RCSP spatial filtering, we used an LDA classifier. The regularization parameters for the RCSP were selected using cross-validation on the training set. More details about the evaluation procedure can be found in [89].

Results

The classification performances obtained on different data sets by the standard CSP and the two RCSP variants mentioned above, namely the SRCSP and the WTRCSP, are reported below in Table 1.4.

Table 1.4: Average classification accuracy (%) obtained with the standard CSP and Regularized CSP variants.

Paper	Data sets	Number of subjects	CSP	SRCSP	WTRCSP
[89]	DSIIaBCICompIV DSIVaBCICompIII DSIIIaBCICompIII	17	75.5	79.2	**79.4**

What can be observed is that both regularized versions of the CSP outperformed the standard, non-regularized, CSP algorithm, leading to increased classification performances. WTRCSP is actually significantly more accurate than the standard CSP ($p < 0.05$, Friedman test).

This approach thus proved effective in making CSP more robust and thus in improving BCI effectiveness. It should be mentioned that other interesting penalty terms have been proposed, in order to deal with known noise sources [93], non-stationarities [94, 95] or to perform simultaneous channel selection

[96, 97]. As we will see in the next section, another approach can be used to robustify CSP.

1.4.2 Robust covariance matrix averaging for CSP

As exposed is Section 1.2, the CSP optimization process is based on the GEVD of the average covariance matrices of each class. Thus, using poorly estimated or noisy average covariance matrices often leads to poor spatial filters, and thus poor BCI performances [88]. Hence, improving average covariance matrix estimators should improve CSP performance. Moreover, covariance matrices are symmetric positive-definite (SPD) matrices which Riemannian geometry can effectively handle [98, 99]. Therefore, in this study, we proposed to robustify CSP by robustly averaging the covariance matrices used in its design, based on Riemannian geometry [100].

Covariance matrix averaging in a Riemannian framework

The standard covariance matrix averaging used in CSP, i.e., the Euclidean average (presented earlier in Equation 1.2) can be very sensitive to outliers. Thus CSP can be made more robust by finding a way of down-weighting noisy trials (e.g., artefactual trials) in the class-covariance estimation. SPD matrices —covariance matrices in our case— belong to a Euclidean space, namely the space of symmetric matrices. For example, a 2×2 SPD matrix A can be written as $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ with $ac - b^2 > 0$, $a > 0$ and $c > 0$. Then symmetric matrices can be represented as points in $\mathbb{R}^3$ and the constraints can be plotted as a cone, inside which SPD matrices lie strictly (see Fig. 1.2). A straightforward approach to average matrices in this space would be to simply use the Euclidean distance δ_e:

$$\delta_e(A, B) = \|A - B\|_{\mathcal{F}}, \tag{1.13}$$

where $\|\cdot\|_{\mathcal{F}}$ denotes the Frobenius norm. The Euclidean geometry of symmetric matrices implies that distances are computed along straight lines according to δ_e (see Fig. 1.2 again).

Implicitly, averaging covariance matrices based on Euclidean geometry corresponds to using formula in Eq. (1.2). This is therefore the standard averaging way used in CSP. However, this Euclidean geometry suffers from several disadvantages. Notably, in addition to the sensitivity to outliers already mentioned, Euclidean geometry suffers from the so-called swelling effect highlighted in [101]. This effect translates the fact that the determinant of the average of two matrices can be bigger than both of their determinants. The implied distortion is then an artifact from the geometry. To avoid this problem, we can

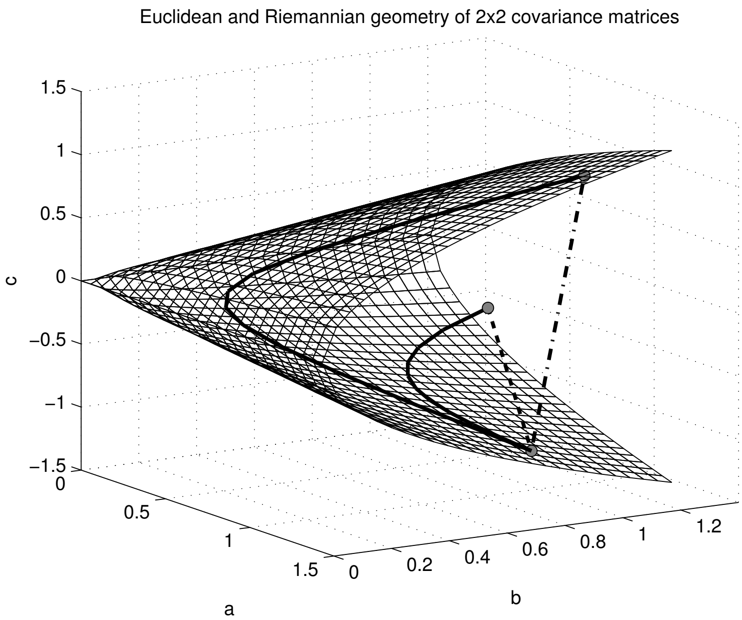


Figure 1.2: Comparison between Euclidean (δ_e - straight dashed lines) and Riemannian (δ_r - curved solid lines) distances measured between points of the space $\mathcal{P}_2$.

use a more natural metric to compare SPD matrices, namely, the *LogEuclidean distance* δ_l [101, 102]:

$$\delta_l(A, B) = \|\log(A) - \log(B)\|_{\mathcal{F}}, \tag{1.14}$$

where $\log(\cdot)$ stands for the matrix logarithm. Finally, as shown in [103], using the proper Riemannian metric, a distance $\delta_r(A, B)$ between two SPD matrices A and B can be computed along curves (namely geodesic - see Fig. 1.2), as follows:

$$\delta_r(A, B) = \|\log(A^{-\frac{1}{2}} B A^{-\frac{1}{2}})\|_{\mathcal{F}}. \tag{1.15}$$

This distance is immune to the swelling effect [101]. It is therefore a good candidate for averaging covariance matrices.

Independently of the chosen distance, the problem of averaging a set of objects in a metric space can be expressed as Eq. (1.16):

$$\min_{\Sigma} \sum_i \delta^2(S_i, \Sigma). \tag{1.16}$$

Using the Euclidean distance δ_e (in Eq. (1.13)), the Euclidean average Σ_E is obtained with the closed-form solution in Eq. (1.2). On the other hand, as shown in [101], when using the LogEuclidean distance δ_l (as in Eq. (1.14)), we have the following closed-form solution:

$$\Sigma_L = \exp\left(\sum_i \log\left(S_i\right)\right), \tag{1.17}$$

where exp(.) denotes the matrix exponential. However, when the Riemannian distance δ_r is used, there is no closed-form solution for computing the Karcher mean Σ_R and optimization techniques [104, 105] are used. In practice, as it was numerically found stable and fast to converge, we use the algorithm proposed in [106].

Even when a suitable distance is used in the Eq. (1.16), the square in the formula makes the problem sensitive to outliers. To remedy this problem, the square is removed from the formula. Then, computing a median is done by solving the following equation:

$$\min_{\Sigma} \sum_i \delta(S_i, \Sigma). \tag{1.18}$$

The iterative algorithm that we used to compute this median has been proposed in [107].

In this work, we therefore studied and compared the use of each of these distances and averaging methods to obtained average covariance matrices for each class in the CSP algorithm.

Evaluation

Altogether, we compared 4 approaches, i.e., the standard Euclidean averaging (Eq. (1.2)) -minimizing δ_e^2- , the LogEuclidean mean -minimizing δ_l^2-, the Karcher mean [106] -minimizing δ_r^2- and the Riemannian median -minimizing δ_r. The resulting average covariance matrices were then used in the CSP optimization. The resulting CSP features were then used as input to a linear SVM. These methods were assessed on 17 subjects from 3 data sets: DSIIaBCICompIV, DSIVaBCICompIII and DSIIIaBCICompIII, already described in Sections 1.3.2 and 1.4.1. See [100] for details.

Results

The classification performances obtained on different data sets by the different covariance matrices averaging methods used with CSP, are reported below in Table 1.5.

What can be observed from these results, is that on average, the Riemannian mean leads to better performances than the standard Euclidean mean with CSP spatial filters, although the performance gain is insignificant. A

Table 1.5: Average classification accuracy (%) obtained with the standard CSP (δ_e^2) and the robust CSP variants based on various covariance matrix averaging methods: Riemannian mean (δ_r^2), Riemannian median (δ_r) and LogEuclidean mean (δ_l^2).

Paper	Data sets	Number of subjects	δ_e^2	δ_r^2	δ_r	δ_l^2
[100]	DSIIaBCICompIV DSIVaBCICompIII DSIIIaBCICompIII	17	78.81	**78.89**	78.48	76.59

more detailed look at the results (see [100] for the detailed results for each subject), reveal than the Riemannian mean actually substantially outperform the Euclidean mean when the the number of EEG channels is small, but not when it is large. Indeed, on data set DSIIaBCICompIV, with 22 EEG channels, the Riemannian mean and median obtained 79.24% and 79.17% classification accuracy on average, versus 76.31% for the Euclidean mean. On Data sets DSIVaBCICompIII and DSIIIaBCICompIII, with 60 and 118 channels, The Euclidean mean outperformed the Riemannian mean with 81.63% of accuracy versus 78.49%. This suggests that Riemannian averaging can indeed make CSP more robust, but only when the dimension is small. A possible explanation for this would be that, as the dimension grew, we reach the limit of the SPD assumption of empirical covariance matrices, and numerical problems appear. In particular, the higher the number of EEG channels, the more likely the signals will be correlated, and thus the more likely some eigenvalues of the covariance matrix will be very close (or equal) to zero. This would thus makes the empirical covariance matrix likely not to be SPD anymore, in which case the Riemannian geometry becomes less efficient as it is designed to handle SPD matrices. Future work would thus be needed to explore dimensionality reduction method in a Riemannian framework, such as [108].

1.4.3 Conclusion on robust spatial filter design

In this section we have shown that BCI can be made more effective by designing spatial filters that are more robust, i.e., spatial filters that can be optimized despite noisy and non-stationary training examples, which EEG signals are. In particular, we have shown that we can regularize the CSP algorithm and have proposed several algorithms to do so, which improved the obtained classification performances on motor imagery BCI data sets. We have also proposed to make CSP spatial filters more robust by improving the averaging of the covariance matrices used in its optimization. We have notably explored Riemannian geometry to do so, which is dedicated to manipulate SPD matrices such as covariance matrices. Our results shown that such approach can also lead to improved classification performances, but only when the number of channel is relatively small.

Altogether this stresses the need for robust methods, here spatial filters, for EEG-based BCI design. Other groups have also experimentally observed this need and proposed alternative methods to robustify CSP, that also proved effective, see, e.g., [93, 109, 94, 95]. The framework and tools proposed here could be further extended by exploring different regularizers or distances. Interestingly enough, we will show in the next section that our RCSP framework can also be used to reduce BCI calibration time.

1.5 Reducing BCI calibration time

As mentioned in the introduction on this manuscript, another point that makes BCI usability poor, is their long calibration times. This is due to the fact that many examples of the user's EEG signals must be recorded in order to calibrate the BCI using machine learning [27]. Unfortunately such calibration and associated necessary data collection are both inconvenient and time consuming. For instance, a typical online BCI based on motor-imagery (i.e., imagination of movements) requires a calibration time of about 20 minutes [110], which is still far too long. As a comparison, nobody would indeed use a computer mouse if it required a 20 minute-long calibration before each use. Therefore, an ideal BCI system should require a calibration time that is as short as possible or even do not require calibration at all.

Therefore, this section explores signal processing tools to reduce or suppress calibration times in oscillatory activity-based BCI [59]. In particular, we proposed a couple of new approaches to reduce calibration time, notably two methods based on artificial EEG data generation and a simple and efficient user-to-user transfer approach based on regularization. We also evaluated a method based on statistical regularization with shrinkage, and methods for subject independent BCI design. They are presented below.

1.5.1 Calibration reduction methods

When studying the standard oscillatory BCI design presented in Section 1.2, it is interesting to note that both the CSP and LDA algorithms require the estimation of covariance matrices. If few training data is available, or if the data available does not reflect most of the variability that can occur during BCI use, the covariance matrices may be poorly estimated and/or not representative of the EEG during use. This would lead to inadequate classifiers or spatial filters. This explains why many examples of EEG signals should be collected in order to calibrate BCI systems, thus making BCI calibration long and tedious. For instance, in the study in [111], the authors found that at least 40 trials per class were necessary to obtain reasonable performances with their motor imagery-based BCI. Such limitations can thus notably be addressed by aiming at obtaining reliable and well estimated covariance matrices from few trials. We present and evaluate below three families of methods to do so: regularizing covariance matrices with shrinkage, user-to-user data transfer and artificial data generation. We also present and evaluate a couple of methods to try to suppress BCI calibration time altogether.

Statistical regularization with shrinkage

Shrinkage consists in using a regularized estimate of the covariance matrices as $\tilde{C}_i = C_i + \lambda I$, where I is the identity matrix. This regularized estimate requires to choose the extent of the regularization with the free parameter λ. Fortunately, a closed-form solution to obtain the optimal value of λ has been proposed, hence avoiding the need for costly procedures such as cross-validation [90]. Incorporating this automatically regularized estimator into CSP (hence leading to an RCSP as described in Section 1.4.1) and LDA algorithms leads to BCI that can be trained with less training data than the standard estimator [112], hence reducing calibration time.

Using data from other subjects: Multi-User BCI

Another approach that can be used to reduce calibration times in BCI is to perform user-to-user transfer, i.e., to transfer data (EEG signals) from users for which many data are available to the target user for which there are few data. In particular, we and others have proposed to use data from other users to perform multi-users covariance matrix estimation [113, 114, 112]. With this approach, the covariance matrix estimate used in CSP and/or LDA for each class can be regularized to resemble the covariance matrices of other users as $\hat{C} = \lambda C + (1 - \lambda)P$ where P can be the average covariance matrix of the other users [113], a weighted average of them [114], or an average of the covariance matrices of selected users [112]. This again leads to an RCSP formulation as proposed in Section 1.4.1. These regularization approaches guide the optimization algorithms towards good solutions, similar to that of the other users, thus enabling a better learning from few data. Here we propose a simple, computationally fast and efficient method to so. Rather than using all other users together as in [113] or weighting or selecting them using heuristics as in [112, 114], we propose to use a sound theoretical framework to assess which other users' data should be used based on their similarity to the data of the target user. In particular, we propose to use Riemannian geometry [98, 99], already mentioned in Section 1.4.2, to measure how different the average covariance matrices of the other users are from that of the target user. Using the Riemannian distance δ_r (see Eq. 1.15), we can easily and quickly identify which other users have covariance matrices that are close to that of the target user and thus use them as regularizers. More precisely, we propose to perform user-to-user transfer by regularizing the target user covariance matrices, for both CSP and LDA covariance matrices, as follows:

$$\hat{C_{target}} = \lambda C_{target} + (1-\lambda)\sum_i \frac{1}{\gamma_i} C_{s_i} \text{ with } \gamma_i = \frac{\delta_r(C_{target}, C_{s_i})}{\sum_j \delta_r(C_{target}, C_{s_j})} \quad (1.19)$$

where C_{target} is a covariance matrix estimated for the target user and C_{s_i} the covariance matrix estimated for the i^{th} other user. In other words, the Riemannian distance enables us to emphasize covariance matrices that are close to that of the target user in the regularization term, and de-emphasize those that are too different. These regularized covariance matrices are then plugged into the CSP (spatial covariance matrices) and LDA (feature covariance matrix) algorithms to perform user-to-user transfer. To avoid the selection of the regularization parameter λ (which is difficult when few training data is available - thus preventing cross-validation use), we re-use the trick introduced by Lu *et al* in [113] and optimize several CSP and LDA pairs, one for each value of λ among of set of possible λ values. When classifying a new trial, these different CSP+LDA pairs are combined by summing the LDA outputs (signed distance of the feature vector from the LDA separating hyperplane) from each of them to determine the final class. This results in a simple, fast, theoretically sound and parameter free multi-user BCI.

Generating artificial trials

The idea behind Artificial Data Generation (ADG) is to generate multiple artificial trials from the few training trials available in order to increase the training set size. Indeed, the problematic need for large amounts of training data is not unique to the BCI field, and can also be encountered in other fields in which machine learning is involved, although the problem might be more severe for BCI. In these fields, some authors proposed to deal with this issue by generating numerous artificial training data from the few original training data available, and use it to augment the training set. This has been shown to lead to increased classification accuracies in fields such as speech processing [115] or hand-writing recognition [116]. We therefore propose here to explore this idea for BCI design. In particular, we propose two ways to generate artificial EEG trials for BCI, by using signal segmentation and recombination in 1) the time domain and 2) the time-frequency domain. Note that both methods should be applied on already band-pass filtered EEG signals.

Signal segmentation and recombination in the time domain: The idea of this first simple approach to ADG is to first divide each training EEG

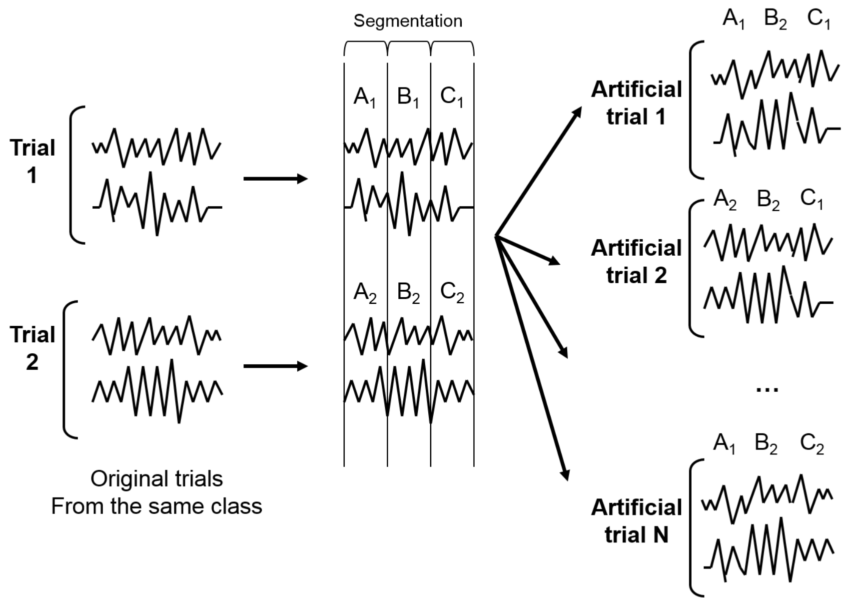


Figure 1.3: Principle of artificial EEG data generation in the time domain

trial into several segments, and then generate new artificial trials as a concatenation of segments coming from different and randomly selected training trials from the same class. More formally, let us denote as $\Omega = \{T_i\}, i \in [1, N]$ as the set of N band-pass filtered EEG trials that are available for training for a given class (note that this ADG is performed separately for each class), $T_i \in \mathbb{R}^{C \times S}$, with S the number of samples in a trial, and C the number of channels. The first step consists in dividing the signals (from each channel) of each training trial T_i into K consecutive and non-overlapping segments $T_i^k \in \mathbb{R}^{C \times S/K}, k \in [1, K]$ (each segment containing the same number of EEG samples). Then, from these segments, we can generate a new artificial trial $\tilde{T}_j$ as $\tilde{T}_j = [T_{R_1}^1 T_{R_2}^2 \dots T_{R_K}^K]$, where $[AB]$ denote the concatenation of the samples from segment A and B (in other words a concatenation of the columns, i.e., along the time dimension), and R_k is a randomly selected integer (random selection with replacement) from $[1, N]$. The whole process is schematized in Figure 1.3. This simple approach enables us to generate a large number of new trials, different from the original ones, but still relevant and likely to be similar to future trials, as they were made from parts of real trials and have the same temporal structure.

Signal segmentation and recombination in the time-frequency domain: While the previous approach is extremely simple, and yet effective (see results section below), it is also brutal. Indeed, simply concatenating segments from different trials may result in inadequate matching between the EEG samples at the boundary between two consecutive segments and thus

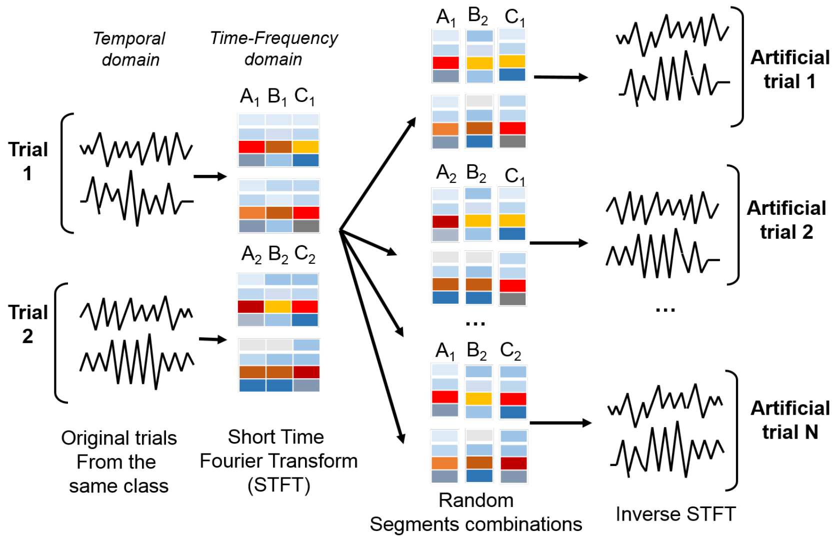


Figure 1.4: Principle of artificial EEG data generation in the time-frequency domain

add some unwanted high frequency noise. To avoid this issue, it could be useful to perform the trial segmentation and random recombinations into the time-frequency domain rather than directly in the time domain. To do so, we first transform each band-pass filtered training trial T_i in a time-frequency representation TF_i by using a Short-Time Fourier Transform (STFT) for each channel. We denote as TF_i^k the k^{th} time window (containing a Fourier spectrum for each channel) of trial T_i in the time-frequency domain. Then, from these time windows, we can generate a new artificial trial in the time-frequency domain $\tilde{TF}_j$ as $\tilde{TF}_j = [TF_{R_1}^1 TF_{R_2}^2 \dots TF_{R_K}^K]$, i.e., by concatenating together STFT windows from different trials from the same class. The final artificial trial $\tilde{T}_j$ is obtained by using inverse STFT on $\tilde{TF}_j$. This process, illustrated in Figure 1.4, is repeated multiple times to generate multiple artificial trials.

Designing subject-independent BCIs

To completely suppress BCI calibration time, it is necessary to build a user-independent BCI system, i.e., to have features and classifiers that work well across users. This is a difficult challenge due to the large between-user variability in oscillatory activity-based BCI. So far, it has been addressed using two main approaches: 1) pooled designs, i.e., calibrating a BCI on the pooled data from multiple users, and 2) ensemble designs, in which user-specific BCIs are combined together to create a user-independent one.

Pooled design A straightforward approach to user-independent BCI design is to optimize CSP filters and the LDA on the combined data from multiple users [117]. We studied this user-independent design with automatic covariance matrix shrinkage (for both CSP and LDA).

Ensemble design More advanced and efficient approaches consist in using ensemble methods. With ensemble methods, one can learn a BCI model (typically CSP+LDA) for each one of the users available, and then combine them to be efficient across users, as in [118]. Here, this combination is achieved by simply using each CSP+LDA obtained for each training user on the new unseen trial, then by concatenating the LDA outputs (signed distance to the LDA hyperplane) together and used them as input to an higher level LDA (also trained previously) which takes the final decision. Automatic covariance matrix shrinkage was used for both CSP and LDA with this approach.

1.5.2 Evaluation

The different calibration reduction or suppression methods described above were evaluated on EEG data from 50 users, from 3 data sets, for different number of training trials, offline. We aimed at finding out how few training data were necessary to achieve a desired level of performance, and thus how long the calibration time would be for different methods.

Data Sets

The methods were evaluated on Data set IIa from BCI competition IV (DSIIaBCICompIV - already described in Section 1.3.2), as well as on two in-house data sets, a workload data set and a mental imagery data set. Both will be described in more details in Chapters 3 and 2 respectively, during the description of the experiments during which they were collected. They are briefly described below.

Workload data set (DSWKL): This data set was recorded while 21 users performed two tasks involving different mental workload levels (easy tasks vs difficult tasks). The cognitive difficulty of the task was manipulated using the N-back task. With this task, users saw a sequence of letters on screen, the letters being displayed one by one, every 2 seconds. For each letter the user had to indicate with a mouse click whether the displayed letter was the same one as the letter displayed N letters before. Each user participated into 12 blocks, alternating between *easy* blocks with the 0-back task (the user had to identify whether the current letter was the letter 'X') and *difficult* block with the 2-back task (the user had to identify whether the current letter was the same letter as the one displayed 2 letters before). Each block contained 30 letter presentations. EEG signals were recorded using 30 EEG channels. This data set is described in more details in [58, 119].

Mental Imagery data set (DSMEI): This data set comprises EEG data from 20 users who performed mental imagery tasks [120, 121]. More precisely, users were prompted to perform either left hand motor imagery, mental rotation of a geometric figure (the figure being displayed on screen) or mental subtractions (successive subtraction of a two digits number from a three digits number, both numbers being displayed on screen). For the purpose of this study, only EEG signals corresponding to left motor imagery and mental rotation of a geometric figure were used.

Comparisons

We studied the performance of each calibration time reduction method and the standard approach (i.e., basic CSP and LDA) for different numbers of training trials. For the calibration time suppression approaches, for each data set, the user-independent BCI were trained on the training EEG trials from all available users expect one, and tested on the testing set of this remaining user. The reported classification performance (percentage of trials whose class was correctly estimated) are those obtained on the testing sets of each user, for which all available trials were used. Further details can be found in [59].

1.5.3 Results

Figure 1.5 display the average classification accuracy (averaged over all users from each data set) obtained by the different methods, on each data set, for different number of training trials. Results first suggest, as expected, that for the standard design, for all data sets, the less trials used for training, the lower the performances. In particular for small training sets, typically with less than 20 trials per class, the performances are very low, near or at chance level (50%), and decrease dramatically when the number of training trials decreases. This confirms the need for numerous training trials for each user, and thus the resulting long calibration time. Then, what can be observed is that ADG approaches often increase classification performances, particularly when few training data is available. These differences are statistically significant for data set DSMEI ($p < 0.01$) for ADG in the time domain and show a trend for ADG in the time-frequency domain ($p = 0.07$). On Data set DSWKL, ADG in the time-frequency domain is significantly better than the baseline ($p < 0.05$). Overall, ADG in the time-frequency domain (ADG-TF) is on average better than ADG in the time-domain (ADG-TD), although not significantly so. Overall, these results indeed support that ADG methods can be used to reduce BCI calibration time, since they can achieve a given performance with less training data than the baseline approach. As could be expected, the User-Independent (UI) methods have lower performances than the user-specific methods when all available training data are used. However, when very few training data

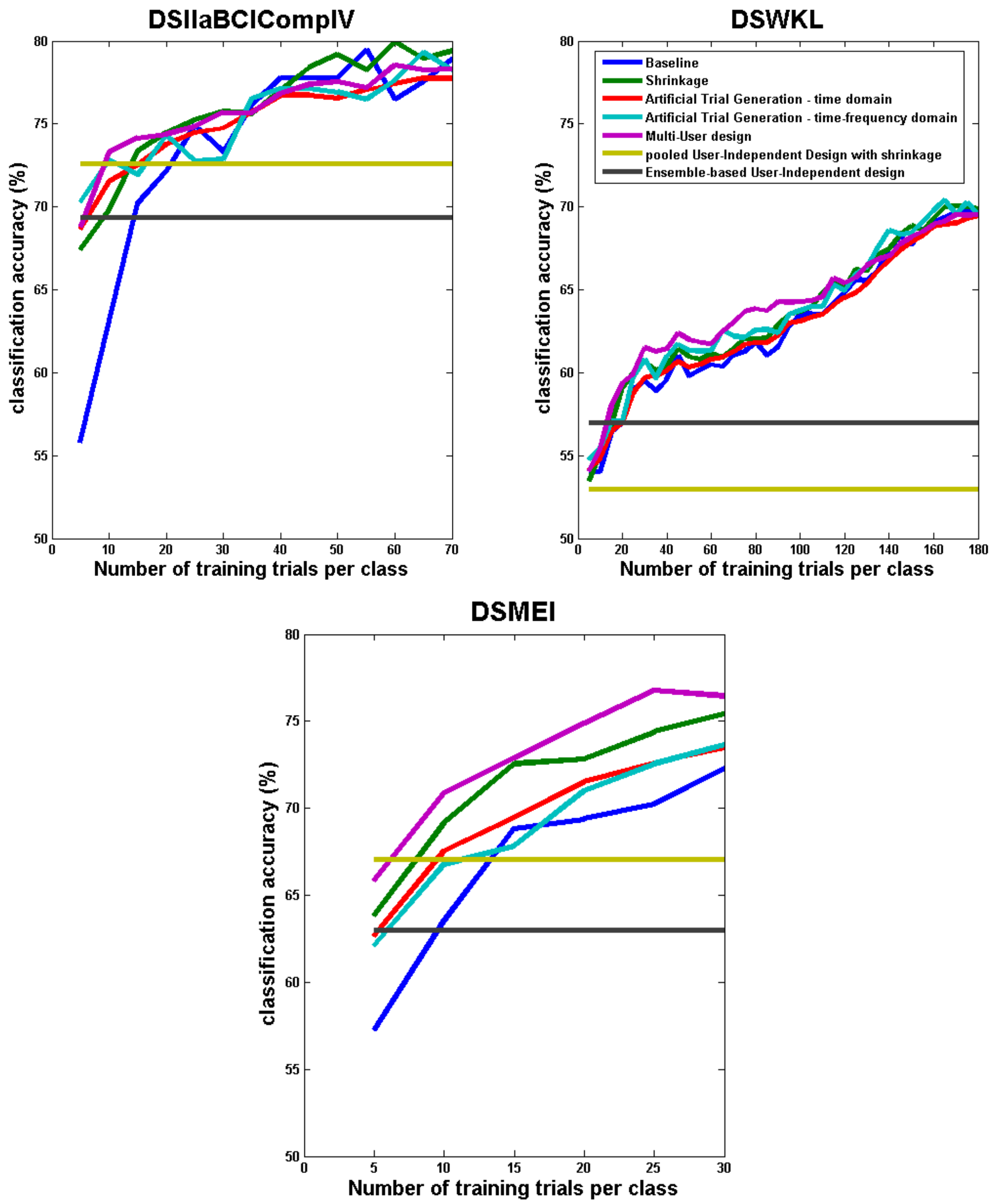


Figure 1.5: Average classification performances (over users) for the different calibration reduction methods and the baseline (standard BCI design - see Section 1.2). Standard deviation bars are omitted for clarity.

are available (i.e., less than 10-15 trials per class), then UI methods often outperform the baseline on average. Results also showed that automatic covariance matrix shrinkage significantly outperformed the baseline ($p << 0.01$), on all data sets. This increase of classification performance is particularly clear when the number of training trials is very low, but is still present even when the maximum number of training trials is used. When comparing ADG-TF and automatic shrinkage, there were no significant differences between the two methods on data set DSIIaBCICompIV and DSWKL, but shrinkage was significantly better than ADG-TF on data set DSMEI ($p < 0.05$). However, combining ADG-TF with shrinkage makes the performance of ADG-TF better. Indeed, ADG-TF+shrinkage outperformed simple shrinkage on data set DSWKL ($p < 0.05$), whereas there was no significant differences between ADG-TF+shrinkage and shrinkage on the two other data sets (see [59]). The proposed Multi-User BCI (MU-BCI) design also proved very efficient. Even when the maximum number of training trials is used, this approach can substantially improve performances, suggesting that it is not only useful for calibration time reduction but also for performance improvement in general. On average over all data sets, MU-BCI notably significantly outperformed all other methods ($p < 0.05$), except shrinkage which was significantly outperformed by MU-BCI only on data set DSMEI ($p < 0.05$). Overall, the best three methods for calibration time reduction are shrinkage, ADG-TF with shrinkage, and MU-BCI. Regarding UI-BCI designs, they are most of the time clearly outperformed by the other approaches, except with only 5 training trials per class. Overall, it is interesting to notice that with only 10 trials per class, several calibration time reduction methods can reach a classification performance equivalent to that of the baseline with 30 to 40 trials per class, hence effectively reducing the calibration time by 3 or 4.

1.5.4 Discussion on Calibration Reduction methods

From the results obtained, we can identify guidelines about which tools to use to reduce or suppress calibration time in which context:

- Automatic covariance matrix shrinkage should always be used, whatever the number of training trials. Indeed, it does not only enable calibration time reduction but also overall performance improvement, even when many training trials are used.
- If data from other users are available, user-to-user transfer is an efficient way to further reduce calibration time or even boost classification performance irrespectively of the number of training data, and should be used as a method of choice. The MU-BCI design proposed in this paper is a fast, simple and efficient method to do so.
- If no data from other users is available, artificial EEG trial generation combined with shrinkage can be used to further reduce calibration time.

– Although user-independent BCI design is possible, performances are still rather poor and need further research to be improved.

There are thus a number of signal processing tools that can significantly and substantially improve classification performance as compared to the standard BCI design when very few training data are available. Good classification performances can thus be maintained with much fewer training data, hence effectively reducing BCI calibration time.

1.6 Discussion and perspective on EEG signal processing

In this chapter, we have proposed contributions at the EEG signal processing level to address the usability issues of BCI, in particular to improve their efficacy and efficiency. First, in order to improve the accuracy of mental command decoding (i.e., BCI efficacy) we have explored alternative features to described EEG signals, namely multifractal cumulants, predictive complexity, waveform length and phase locking values. We have also proposed optimal spatial filters to maximize the discriminative power of the last two features. We have shown that our optimal spatial filters indeed significantly and substantially improved the classification accuracy obtained with these features. Our results also shown that combining these alternative features with the classical features used for BCI, i.e., band power features (possibly with CSP spatial filtering) led to increased classification accuracy, and thus to more effective BCI. Still to improve the efficacy of BCI, we have also designed more robust CSP spatial filters, that can be better optimized despite noise and non-stationarities. We have proposed a regularization framework and algorithms to do so, as well as robust averaging techniques based on Riemannian geometry. Both methods led to improved classification accuracies on average. Finally, we have proposed signal processing tools to improve BCI efficiency by reducing their calibration times. To do so, we have proposed methods based on automatic covariance matrix shrinkage, artificial EEG data generation, and user-to-user transfer. Our evaluations suggested that these algorithms could reduce calibration by a factor 3, as they need about 10 training trials per class to reach the same accuracy as the standard design with 30 training trials per class. Overall these methods contributed to increase BCI accuracy, robustness and to shorten their calibration time.

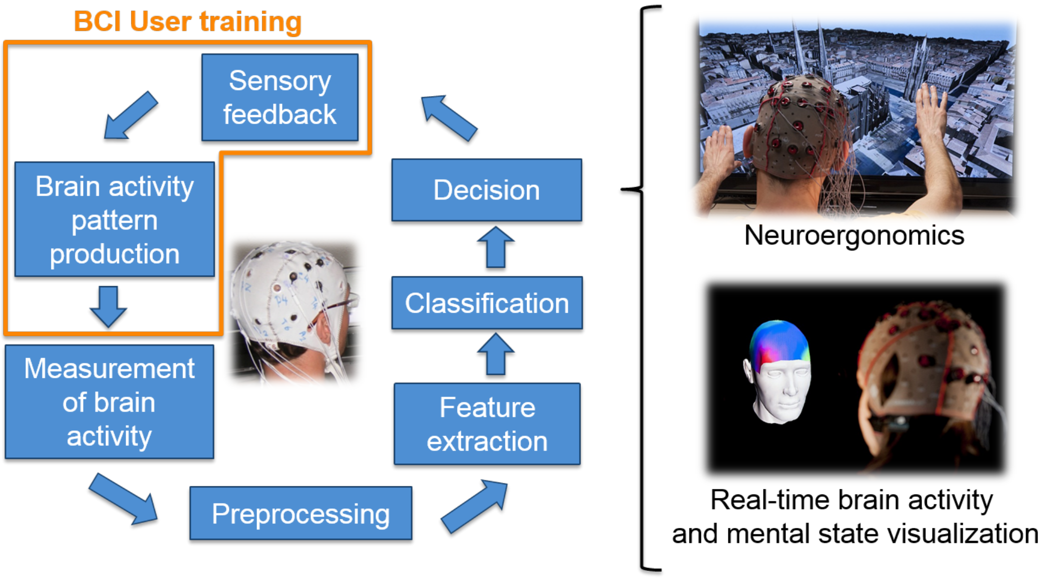
However, despite these improvements and other brought by the research community, there is still a lot to improve in EEG signal processing for BCI. Indeed, the obtained accuracies are still far from being satisfactory for practical applications, as the rate of erroneous commands recognized by the BCI is still substantial. The calibration times are also too long, since there is still a need for a calibration: ideally, the BCI should be usable immediately by any new user, like most other interaction devices. Thus more research on EEG signal processing is still needed. In order to improve BCI accuracy, one could explore additional alternative features, how to combine them optimally, and design spatial filters that are optimal for each of these new alternative features. To make BCI robust to noise and non-stationarity, it would be necessary to study from a fundamental point of view what are their precise causes and neurophysiological signatures. In other words, it would be necessary to identify the sources of variance of EEG signals, in order to design algorithms that are invariant to these sources. To go even further, we should identify the sources of variance across time, contexts and users, to make BCI invariant to all these variations. This should drastically improve BCI performances, and suppress the need for calibration altogether.

In order to obtain new EEG signals representations and invariance properties, an interesting approach to further explore is EEG signal processing based on Riemannian geometry, which we presented before. Indeed, the Riemannian distance is affine invariant and can, as exposed before, be used to directly manipulate covariance matrices. These properties recently enable Riemannian geometry-based EEG classification to show their superiority to other classical EEG signal processing approaches based on feature vector classification, by being the winning methods on a couple of recent brain signal classification competitions, notably the "'DecMEG2014"' (https://www.kaggle.com/c/decoding-the-human-brain) and the "'BCI challenge 2015"' (http://neuro.embs.org/2015/bci-challenge/) competitions. Moreover, the methods used in these designs were rather simple, based on band-pass filtered EEG signal covariance matrices, i.e., reflecting the EEG signals band power. Other EEG signals representations as SPD matrices could be explored and used to further improve BCI performances. The traditional tools used for EEG signal processing could be redesigned in a Riemannian framework to benefit from its properties. Some of them have already been designed and used for BCI, such as artificial data generation [122], dimensionality reduction [108] or metric learning [123]. It would be promising to design and explore further Riemannian tools for BCI such as supervised dimensionality reduction, multi-task learning or feature tracking, among other. The interested reader can refer to our recent paper in [124] as well as to [99] for more information on this promising recent research direction.

Finally, making EEG-based BCI fully usable cannot be done by working only at the EEG signal processing level. Indeed, as the name brain-computer interaction suggests, not only the computer is important but also the brain, i.e., the BCI user. In particular, if the user cannot produce stable and distinct brain activity patterns, even the best signal processing algorithm in the world will fail at recognizing them. In this case the BCI can be neither efficient nor effective. In the next chapter, we thus propose contributions at the BCI user level, to ensure those users can acquire high quality BCI control skills.

Chapter 2

Understanding and improving BCI user training



Selected related Publications:

- C. Jeunet, E. Jahanpour, F. Lotte, *Why Standard Brain-Computer Interface (BCI) Training Protocols Should be Changed: An Experimental Study*, Journal of Neural Engineering, vol. 13, no. 3, 036024, 2016
- C. Jeunet, B. N'Kaoua, F. Lotte, *Advances in User-Training for Mental-Imagery Based BCI Control: Psychological and Cognitive Factors and their Neural Correlates*, Progress in Brain Research, 2016
- C. Jeunet, B. N'Kaoua, S. Subramanian, M. Hachet, F. Lotte, *Predicting Mental Imagery-Based BCI Performance from Personality, Cognitive Profile and Neurophysiological Patterns*, PLOS ONE, vol. 10, no. 12, e0143962, 2015

- J. Schumacher, C. Jeunet, F. Lotte, *Towards Explanatory Feedback for User Training in Brain-Computer Interfaces*, IEEE International Conference on Systems Man and Cybernetics (IEEE SMC), pp. 3169-3174, 2015
- C. Jeunet, C. Vi, D. Spelmezan, B. N'Kaoua, F. Lotte, S. Subramanian, *Continuous Tactile Feedback for Motor-Imagery based Brain-Computer Interaction in a Multitasking Context*, Human-Computer Interaction - INTERACT, pp 488-505, 2015
- F. Lotte, C. Jeunet, *Towards Improved BCI based on Human Learning Principles*, 3rd International Winter Conference on Brain-Computer Interfaces, invited paper, pp. 37-40, 2015
- J.E. Huggins, C. Guger, B. Allison, C.W. Anderson, A. Batista, A.-M. Brouwer, C. Brunner, R. Chavarriaga, M. Fried-Oken, A. Gunduz, D. Gupta, A. Kübler, R. Leeb, F. Lotte, L.E. Miller, G. Müller-Putz, T. Rutkowski, M. Tangermann, D.E. Thompson, *Workshops of the Fifth International Brain-Computer Interface Meeting: Defining the Future*, Brain-Computer Interfaces, vol. 1, no. 1, pp. 27-49, 2014
- F. Lotte, F. Larrue, C. Mühl, *Flaws in current human training protocols for spontaneous Brain-Computer Interfaces: lessons learned from instructional design*, Frontiers in Human Neurosciences, vol 7., no. 568, 2013
- L. Bonnet, F. Lotte, A. Lécuyer, *Two Brains, One Game: Design and Evaluation of a Multi-User BCI Video Game Based on Motor Imagery*, IEEE Transactions on Computational Intelligence and AI in Games (IEEE T-CIAIG), vol. 5, num. 2, pp. 185-198, 2013

Scientists that I (co-)supervised for this work:

- PhD student:
  - Camille Jeunet
  - Jelena Mladenovic
- Engineers:
  - Alison Cellard
  - Boris Mansencal
- Master students:
  - Julia Schumacher
  - Loic Renault
  - Emilie Jahanpour
  - Léa Pillette
  - Suzy Teillet
  - Benjamin Muzart
  - Manon Bonnet-Save

Related research projects:

- Inria Project Lab BCI-LIFT, 2015-2018 (Work Package Leader)
- ANR Project REBEL, 2016-2019 (Principal Investigator)

2.1 Introduction

To operate a BCIs, users have to encode commands in their EEG signals, typically using mental imagery tasks leading to specific EEG patterns. The machine has to decode these patterns by using signal processing and machine learning. So far, most research efforts - including ours - have addressed the usability issue of BCIs by focusing on command decoding only [53, 26, 27, 40]. While this has contributed to increased performances (see, e.g., the previous chapter), correct mental command decoding rates are still relatively low and BCI illiteracy/deficiency still high [53, 2]. Thus, the reliability issue of BCI is unlikely to be solved by solely focusing on command decoding.

BCI control is known to be a skill that must be learned and mastered by the user [28, 125]. Indeed, a BCI user performance (i.e., how accurately his/her mental commands are decoded) become better with practice and BCI training leads to a re-organization of brain networks as with any motor or cognitive training [28, 125]. Therefore, to ensure a reliable BCI, users must learn to successfully encode mental commands in EEG signals, with high signal-to-noise ratio. In other words they should be trained to produce EEG patterns that are as stable, clear and distinct as possible. With poor BCI command encoding skills, even the best signal processing algorithms will not be able to decode commands correctly. Unfortunately, how to train users to encode these commands has been rather scarcely studied so far. As a consequence, the best way to train users to successfully encode BCI commands is still unknown [53, 2, 28]. Worse, as we argue in this chapter, current user training approach in BCI are actually even inappropriate, and most likely a major cause of poor BCI performance, and high BCI deficiency rates. Therefore, in this chapter, we present our theoretical, experimental and methodological work on BCI user training, aiming at 1) understanding the limitations and properties of current BCI user training methods and 2) proposing new user training approaches to improve this training towards ensuring BCI users can acquire high quality control skills.

This chapter is organized as follows: First, Section 2.2 will present our work dedicated to identify the many limitations of current user training protocols, both at the theoretical and practical levels. Then, in order to improve these limitations, Section 2.3 will present our work towards understanding the impact of the user's profile on BCI performances, with the longer term objective to design adapted and adaptive BCI training. Section 2.4 will describe our contributions towards designing better feedback for BCI user training. Finally, Section 2.5 proposes some discussions and perspectives on this line of research. It should be noted that a substantial part of this chapter is based on the work conducted by Camille Jeunet for her PhD thesis that I am co-supervising.

2.2 Limitations of current BCI user training approaches

As mentioned earlier, a major cause of poor BCI usability may be inappropriate user training approaches. We therefore studied standard Mental Imagery (MI)-BCI training approaches, both from a theoretical and practical point of view, with the goal of identifying their possible limitations.

2.2.1 Standard BCI user training protocol

Training users to acquire BCI control skills is usually performed using Neurofeedback [1] [28]. BCI neurofeedback training principles mostly depend on the type of BCI category used [1]. With the *Operant Conditioning (OC)* approach, the EEG signal decoder (classifier) is fixed and unknown to the user, and this user has to find out how to control a cursor by modulating his/her brain activity in a specific way. Using this kind of approach, the training can last for weeks or even months before the user can control the BCI. This was the approach used to successfully design the first BCI systems [30, 5]. With the *Machine Learning (ML)* approach, the EEG decoder (classifier) is optimized on examples of EEG signals collected from the user while he/she performs the targeted mental tasks. With this approach the training time before the user can control the BCI is much shorter (about 20 minutes for 2 classes), see, e.g., [126, 18]. This is the most used approach.

A typical MI-BCI user training protocol, based on the ML approach, is the one proposed by the Graz group [16] which is a widely used BCI training protocol [28]. Most other existing MI-BCI training protocols can be seen as variants of the Graz training protocol as they use similar timings, feedback and training tasks, see, e.g., [126, 18]. The most used tasks in the context of the Graz protocol are motor-imagery tasks (such as the imagination of hand movements) which are known to be associated with an activation of the motor cortex. The Graz protocol is divided into two steps: (1) training of the system and (2) training of the user. During the first step, the user is instructed to perform several successive motor imagery tasks such as the imagination of left- and right-hand movements. From the recorded EEG signals, the system extracts characteristic EEG patterns which are specific to each MI task. These extracted patterns are then used to train a classifier to recognize these tasks. Step 2 consists in training the user. To do so, the user is instructed to perform the same MI tasks, but this time feedback (provided by the classifier, which was optimized in Step 1) is provided to inform the user which MI-task the system has recognized. Thus, the goal of the user will be to find strategies so

1. Neurofeedback consists in providing the user with a real-time feedback about his/her own brain activity so that he/she can learn to voluntarily control it.

that the system recognizes the mental task he/she is performing. This training protocol is most often performed over different sessions divided into runs, one session typically including 4 to 6 runs, in order to avoid the fatigue which is usually felt after more runs. Runs are themselves divided into trials, usually between 10 to 20 per class (i.e., per MI-task). One trial typically lasts 8s. At the beginning of each trial, a fixation cross is displayed to announce the start of the trial and to avoid eye movements during the following 2-second long rest period. Then, after 2s, a beep is used to trigger the attention of the user and prepare him/her for the oncoming instruction. One second later, at t = 3s, the instruction appears as an arrow the direction of which indicates the MI task to be performed, i.e., an arrow pointing left indicates a left hand MI and an arrow pointing right a right hand MI. From t = 3.250s, a feedback is provided for 4s in the shape of a bar the direction of which indicates the mental task that has been recognized and the length of which represents the confidence of the system in the recognition of the MI-task. This sequence of events is depicted in Figure 2.1.

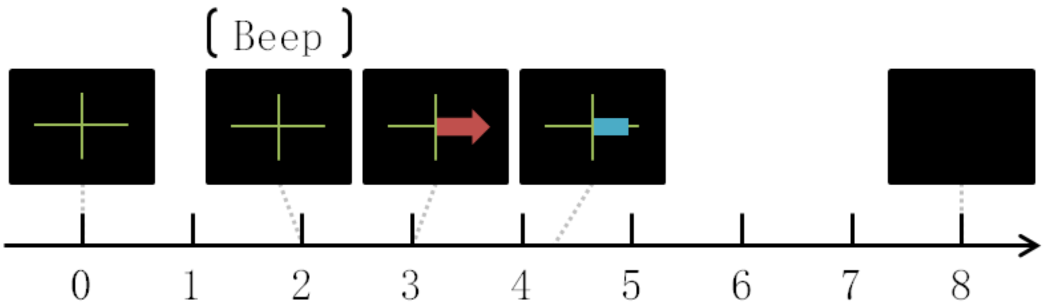


Figure 2.1: Timing of one trial in the Graz Protocol.

While different BCI users training protocols have been proposed in the literature, most of them are based on the Graz protocol described here-above or are similar to it. It is therefore necessary to question whether such a training protocol is indeed the most appropriate for BCI user training.

2.2.2 Theoretical limitations

Is the type of training approach mentioned above really the best way we can train our users to gain BCI control? To answer that question, we have studied the literature from the fields of human learning, educational science and instructional design [127, 60]. These fields have indeed studied across multiple disciplines, e.g., language learning, motor learning or mathematical learning, what are the principles and guidelines that can ensure efficient and effective training approaches. We have then compared such principles and guidelines to the training approaches currently used in BCI. In short, we have shown that standard BCI user training approaches do not satisfy general human learning and education principles ensuring successful learning [128, 129, 130, 131]. Notably, typical BCI feedback is corrective only, i.e., it only indicates users whether they performed the mental tasks correctly. Oppositely, human learning principles recommend to provide explanatory feedback by indicating what was right or wrong about the task performed [129, 130]. BCI feedback is also usually unimodal, based only on the visual modality, whereas exploiting multimodality is also known to favor learning [128, 132, 133]. Moreover, training tasks should be varied and adapted to the user's skills, traits (personality or cognitive profile) and states [130, 129, 131]. They should also include self-paced training tasks, to let the user explore the skills he/she has learned. BCI training tasks, in contrast, are fixed over time and users, synchronous and repeated identically during training. The positive role of the social context and the social interactions to which the learner participates has also been stressed and should be exploited to maximize learning efficiency and task performance [134, 135, 136]. So far, the social context in BCI training is usually not considered at all. Finally and intuitively, training environments should be motivating and engaging, whereas current BCI training environments essentially ignore the user motivation, resulting in plain and boring training. Unfortunately, there are many other training principles and guidelines not satisfied by classical BCI training (see [127, 60] for more details).

2.2.3 Practical limitations

While instructive, the study above only provided theoretical considerations about training approaches. It is therefore necessary to concretely assess whether training approaches used in BCI are appropriate to train a skill. Moreover, it is necessary to perform this evaluation independently of BCI, to rule out possible biases due to BCI complexity, non-stationarity and poor signal-to-noise ratio. Indeed, if a BCI training results in poor performances (i.e., the subject fails to obtain BCI control), this might not be due to the training protocol itself but simply to poor EEG signal processing, noisy or non-stationary signals, or to the fact that the relevant neural signals cannot be found in the EEG signals of the user due to the orientation of the user's cortex, for instance. Thus here, we propose to study these BCI training approaches without using a BCI: participants were asked to learn specific and simple motor tasks using the same feedback and training tasks used for MI-BCI [137]. We then studied how well they could learn such motor tasks to assess the quality of the training approach, independently of BCI use.

Method

Participants were asked to learn to perform two motor tasks: drawing triangles and circles with a pen on a graphic tablet (see Figure 2.2), using standard MI-BCI training approaches [16]. Indeed, as with MI-BCI, in which users have to learn a suitable MI strategy, the participants here had to learn the strategy which allows the computer to correctly recognize their drawing, e.g., they had to identify the suitable shape size, angles or speed of drawing.

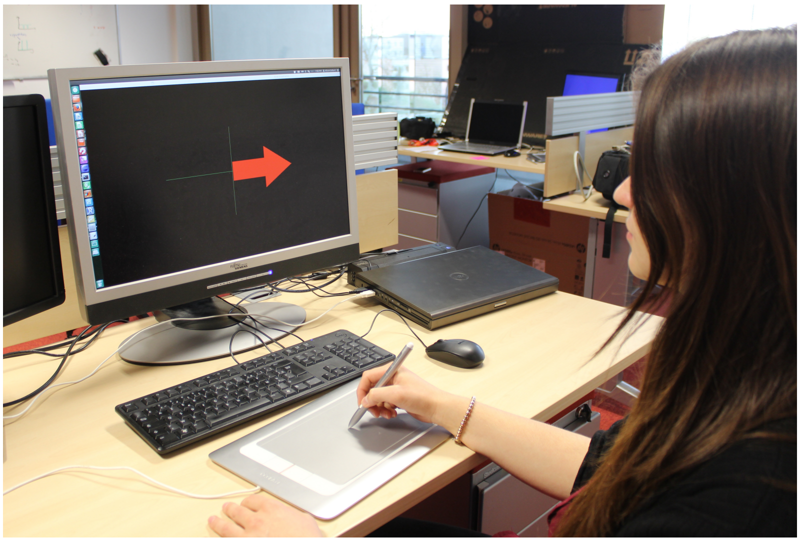


Figure 2.2: Experimental setup to train our users to perform simple drawing tasks using the same feedback and protocol as for MI-BCI.

54 BCI-naive and healthy participants (20 females; aged 25.1 ± 4.6 year-old) took part in this study were they had to learn to draw circles and triangles that can be recognized by the computer during different runs. Each run comprised several trials, whose structure and timing was exactly the same as in the standard Graz BCI training protocol described in Section 2.2.1. A left arrow indicated here to draw a circle, and a right arrow to draw a triangle. The feedback was also exactly the same, i.e., a blue feedback bar whose direction indicated the shape recognized by the classifier (left: circle, right: triangle) and its length was proportional to the classifier output (i.e., the distance to the classifier separating hyperplane), as with MI-BCI. Each participant participated to 4 such runs. They were provided with the following instruction: “Your goal is to find the right strategy so that the system recognizes as well as possible the shape you are drawing, which will concretely correspond to having the feedback bar as long as possible in the correct direction: left for circles and right for triangles”.

In order to discriminate triangular from circular pen movements on the graphic tablet, we used a pattern recognition approach as in BCIs, with an LDA classifier (see [137] for details). This LDA classifier - a subject independent one trained on the data from 2 persons (1 left-handed, 1 right-handed) - could discriminate triangles from circles with 73.8% classification accuracy (10-fold cross-validation on the training set), which is an accuracy equivalent to the average accuracy of an MI-BCI [50]. The output of the LDA was mapped to the direction and length of the feedback bar, as in a typical MI-BCI.

Results

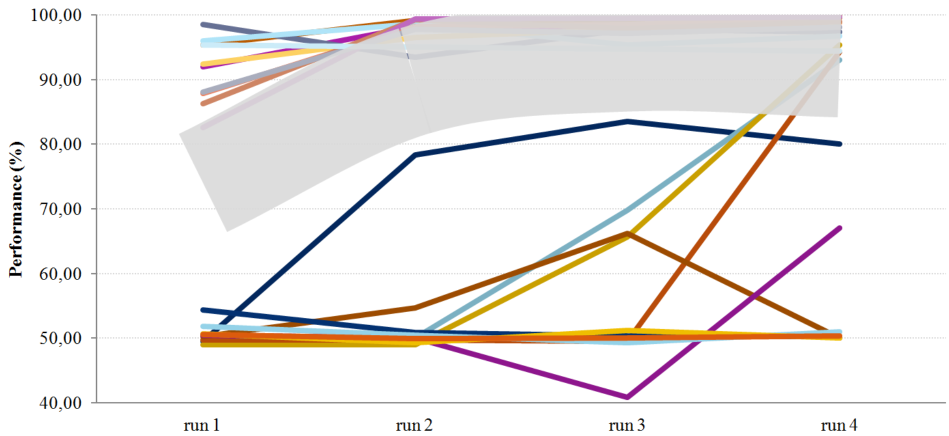


Figure 2.3: Graphic representing the performance of the participants (mean classification accuracy) as a function of the run. We chose to represent the 10 best and 10 worst performers. The average performance of the 34 other participants is represented by the large grey line.

To study how well subjects could learn the motor tasks, we measured their performance as the average classification accuracy obtained to discriminate triangular from circular pen movements, averaged over the whole feedback period. The classification accuracy obtained for each run by each participant is displayed on Figure 2.3. Results showed that 45 out of 54 participants managed to learn the task, i.e. obtained more than 70% average performance [2] -classification accuracy- [138] ($\bar{X} = 89.09\%$; $SD = 6.35$; $range = [72.84, 98.26]$). However 9 participants (i.e., 16.67% of them) failed to learn how to perform correctly such simple gestures ($\bar{X} = 55.68\%$; $SD = 6.35$; $range = [50.23, 65.64]$). Thus, one can hypothesize that BCI illiteracy could not only be due to the user, but also partly to the training protocol. Indeed, it has been hypothesized that BCI illiteracy/deficiency could be due to the user, who may

2. This 70% accuracy is a threshold often used in the BCI community to distinguish subjects that achieved BCI control from those who did not achieve such a control [53]

fail to produce the desired EEG patterns [53]. Our experiment suggests that some subjects may fail to reach BCI control because the training protocol is not suited to everyone, and does not favor learning in general. The interested reader can refer to [137] for more details on this experiment.

2.2.4 Conclusion on the limitations of the current BCI user training approaches

From the studies presented above, it seems that theory (notably from educational science) and practice (motor task training with feedback used in BCI) both suggest that current BCI training approaches are most likely highly suboptimal. This a bad news for current BCI systems and users, but a good news for BCI research and BCI future: it means there is a lot of room for improvement, and many exciting and promising research directions to explore. In particular, this suggest BCI could be made much more usable by making their user training approaches satisfy principles and guidelines from educational science. This notably means that BCI training tasks should be made adaptive and adapted to each user. To do so, we should find out about how the user's profile impact BCI learning and performances. This is what Andrea Kübler's group in Wurzburg has started to do [139, 136] and that we are complementing with the studies presented in the next Section. Satisfying human learning principles also means exploiting multimodal feedback and providing a feedback that is explanatory rather than purely corrective. Our works presented in Section 2.4 go in that direction.

2.3 Understanding the impact of the user's profile on BCI performance

With the longer term objective of providing adaptive and adapted training to our BCI users, we first had to identify the impact of the user profile on BCI performances. The idea is to find out which types of users, notably in terms of psychological or cognitive factors, can successfully control a BCI or not, and why. We conducted a couple of experiments to do so. First we trained several people to control a 3-class Mental Imagery based BCI (involving motor and non-motor mental tasks) over several days, and administered them several psychological questionnaires to identify their profile. We then looked for correlations between this profile and their BCI performances. This study is described in Section 2.3.1. It revealed a number of relevant factors that are related to BCI performances. Thus, we conducted another study in which we assessed such factors in a purely motor imagery-based BCI, which confirmed one of them. This is described in Section 2.3.2. Finally, based on these results and on an survey of the literature, we proposed an overview of the factors influencing BCI performances that should be taken into account for an adaptive and adapted BCI training, in Section 2.3.3.

2.3.1 Impact of the personality and cognitive profile on Mental Imagery-based BCI performances

Our first study aimed at investigating whether there were relationships between the user's profile and his/her BCI performances with a mental-imagery based BCI [120]. We choose not to study a pure motor imagery-based BCI as there were already some works [140, 141, 142, 139, 143] on this topic. Moreover, we performed this study with several sessions of user training since most existing studies were performed on a single day of training [50, 139, 140, 141]. We therefore trained users to control a 3-class mental imagery-based BCI while measuring their profiles with questionnaires, as described here after.

Protocol

18 BCI-naive participants (9 females; aged 21.5 ± 1.2) took part in this study. Each participant took part in 6 sessions, on 6 different days spread out over several weeks. The three MI-tasks that participants had to learn to perform were left-hand motor imagery, mental rotation and mental subtraction. They were chosen according to Friedrich et al. [17], who showed that these tasks were associated with the best performance. "Left-hand motor imagery" (*L-HAND*) refers to the kinesthetic continuous imagination of a left-hand movement, chosen by the participant, without any actual movement [17]. "Mental rotation" (*ROTATION*) and "mental subtraction" (*SUBTRACTION*) correspond respectively to the mental visualization of a 3 Dimensional shape rotating in a 3 Dimensional space [17] and to successive subtractions of a 3-digit number by a 2-digit number (ranging between 11 and 19), both being randomly generated and displayed on a screen [17].

During each run, participants had to perform 45 trials (15 trials per task), each trial lasting 8s (see Fig. 2.4). At t=0s, an arrow was displayed with a left hand pictogram on its left (*L-HAND* task), the subtraction to be performed on top (*SUBTRACTION* task) and a 3D shape on its right (*ROTATION* task). At t=2s, a "beep" announced the coming instruction and one second later, at t=3s, a red arrow was displayed for 1.250s. The direction of the arrow informed the participant which task to perform, e.g., an arrow pointing to the left meant the user had to perform a *L-HAND* task. Finally, at t=4.250s, for 4s, a visual feedback was provided in the shape of a blue bar, the length of which varied according to the classifier output. Only positive feedback was displayed, i.e., the feedback was provided only when there was a match between the instruction and the recognized task, as in [17]. During the first run of the first session (i.e., the calibration run), as the classifier was not yet trained to recognize the mental tasks being performed by the user, it could not provide a consistent feedback. In order to limit biases with the other runs, e.g., EEG changes due to different visual processing between runs, the user was provided with an equivalent sham feedback, i.e., a blue bar randomly appearing and varying in length, and not updated according to the classifier output as in [17].

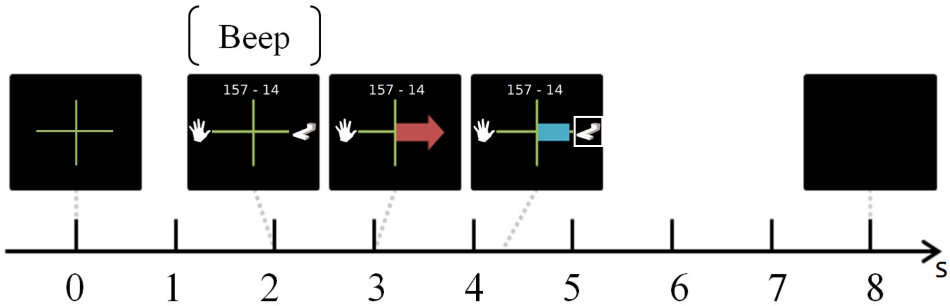


Figure 2.4: Timing of a trial for the 3-class mental imagery based BCI training.

In terms of EEG signal processing, the BCI was designed around a classical CSP+LDA pipeline, with 3 sets of CSP filters (one set for each class versus the other two classes) and a multi-class shrinkage LDA with a one-versus-the-rest scheme, see [120] for details. The CSP and LDA classifiers were calibrated on the EEG data collected in the first run of the first session. To reduce between session variability, the LDA classifiers' biases were re-calculated after the first run of the sessions 2 to 6, based on the data from this first run, as in [17].

At the beginning of each of the 6 sessions, participants were asked to complete different validated psychometric questionnaires, to assess different aspects of their personality and cognitive profile, that have been related to learning in the literature. These questionnaires were the *Wechsler Adult Intelligence Scale* (WAIS-IV) [144], assessing different Intelligent Quotient (IQ) dimensions, the *Corsi Block task* [145] measuring visuo-spatial short term and working memory abilities, the *Revised Visual retention test* [146] quantifying visual retention abilities as well as perceptive organization, the *Learning Style Inventory* (LSI) [147] identifying the students' preferred learning styles, the *16 Personality Factors - 5* (16 PF-5) [148] measuring sixteen primary factors of personality as well as five global factors of personality, the *Internal, Powerful others and Chance scale* (IPC) [149] assessing the locus of control, the *State Trait Anxiety Inventory* (STAI) [150] measuring the stress state and trait, the *Bruininks-Oseretsky Test of Motor Proficiency* (BOT-2) [151] evaluating motor abilities and the *Mental Rotation test* [152] measuring spatial abilities.

Results

Eighteen participants took part in this experiment. The data of one outlier participant were rejected since, with a mean performance of 67.21%, he outperformed (by more than two SDs) the group's mean performance over the six sessions ($\bar{X}_{group} = 52.50\%$; $SD = 5.62$). Thus, the following analyzes were based on the data of 17 subjects.

Bivariate Pearson correlation analyzes revealed correlations between MI-BCI performance and (1) Mental Rotation scores [r = 0.696, $p < 0.005$], (2) Tension [r = -0.569, $p < 0.05$], (3) Abstractness ability [r = 0.526, $p < 0.05$] and (4) Self-Reliance [r = 0.514, $p < 0.05$]. Tension, abstractness and self-reliance were assessed by the 16 PF-5. High "tension" scores reflect highly tense, impatient and frustrated personalities. The Self-Reliant trait, also called self-sufficiency, reflects the learners' ability to learn by themselves, i.e., in an autonomous way. Finally, abstractness refers to creativity and imagination abilities. Among these four factors, only the Mental Rotation score reached significance after the Positive False Discovery Rate correction for multiple comparisons [$p < 0.05$] [153].

A Step-Wise Linear Regression was used in order to determine a predictive model of each user's average MI-BCI performance obtained across the different training sessions. This regression resulted in a first model, called MODEL ♯1, including six factors [R^2_{adj} = 0.962, p < 0.001]: Mental Rotation score, Self-Reliance, Memory Span, Tension, Apprehension and the "Visual/Verbal" subscale of Learning Style . MODEL ♯1 explains more than 96% of the performance variance of the data set.

In order to evaluate (1) the stability and (2) reliability of MODEL ♯1, step-wise linear regressions were then performed using a leave-one-subject-out cross validation process. During the *first step*, 17 models were generated, each of them based on the data of all the participants except one (i.e., the training data set). This *first step* allowed to assess the *stability* of the model by comparing the factors included in each of the models to the ones included in MODEL ♯1. During the *second step*, each of these models was tested on the only participant not included in the respective training data sets (i.e., the testing data set). This *second step* aimed at determining the *reliability* of the models. Each model generated from the training data set enabled to determine a predicted performance as well as a confidence interval for the corresponding testing data set.

The first step of the leave-one-subject-out cross validation process revealed the instability of MODEL ♯1. Indeed, only 5 out of 17 models included the same factors as MODEL ♯1. In 11 out of 17 models, 2 or more factors were different from MODEL ♯1. Results of the second step revealed that the real performance of 9 out of 17 participants fell within the predicted confidence interval, with an absolute mean error ($Perf_{predicted}$ - $Perf_{real}$) of 2.68 points

(*SD* = 2.37, *range*: [0.38, 8.98]).

In Model ♯1, the "mental rotation" factor was selected first in the regression and highly correlated with performance (r=0.696), which demonstrates its strong influence on the model. While being consistent with the nature of the tasks performed by the participants, this strong influence was likely to hide the effect of other important factors [154, 155]. Consequently, a second regression analysis was performed without the mental rotation variable. It resulted in a model, called Model ♯2 [R^2_{adj}=0.809, p < 0.001] which includes 4 parameters: Tension, Abstractness, the Learning Style "Active/Reflective" subscale and Self-Reliance. As was done for Model ♯1, the stability and reliability of Model ♯2 were assessed using a leave-one-subject-out cross validation process. Results revealed that among the 17 models generated, 10 included exactly the same factors as the ones included in Model ♯2: Tension, Abstractness, the "Active/Reflective" Learning Style subscale and Self-Reliance. In 5 out of the 7 remaining models, only one factor, Self-Reliance, was missing. The real performance of 14 out of 17 participants fell within the confidence interval. In order to ensure that the successful prediction of BCI performance using the personality and cognitive profiles of the users was not due to chance, a permutation test was performed (see [120] for details). It indicated that the mean absolute error of 2.87 that we obtained was better than chance with p < 0.01. This means our model can indeed generalize to new subjects and predict their MI-BCI performances from their personality and cognitive profile.

Fig. 2.5 outlines women's results on top and men's results on the bottom. Graphs on the left represent each participant's real (left) and predicted (right) performance for each participant, with the corresponding confidence intervals. Graphs on the right represent the Mental Rotation scores for all the participants. Women and men were separated due to the important gender effect associated with this test [152]. Women's mean score is 19.13/40 (*SD*: 6.29, *range*: [5, 27]). Men's mean score is 29/40 (*SD*:6.56, *range*: [18, 35]). Women's and men's mean scores are represented as a horizontal line on the graphs on the right of Fig. 2.5. The rectangle surrounding this line represents the mean ± 1SD interval. Only 3 participants, one woman and two men, are below this interval: S14, S23 and S28. It is noticeable that the same participants, i.e. S14, S23 and S28, (1) had lower real MI-BCI performance than the one predicted by the model and (2) had lower mental rotation scores than the average.

Discussion and Conclusion

In this work, we proposed a predictive model of MI-BCI performance based on the data of 17 participants. The important number of runs (30, spread over 6 sessions) attenuated the between-session variability (which could be due, e.g., to fatigue or motivation fluctuation, cap position variation, etc.) and

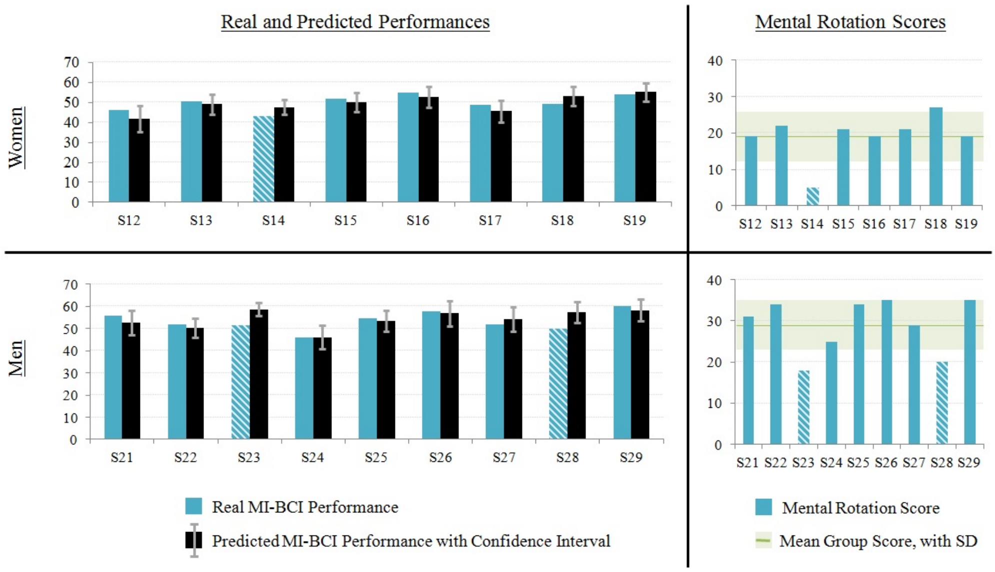


Figure 2.5: Real and predicted BCI performance as well as Mental Rotation scores according to the gender. Women's results are shown at the top, men's results on bottom. On the left, the graphical representation of the real (left) and predicted (right) BCI-performance of each participant, with the corresponding confidence intervals. On the right, the mental rotation scores of each participant with the horizontal line representing the mean score of the group. The three participants for whom the model overrated the performance are those with the lowest mental rotation scores (striped participants).

thus enabled to more precisely estimate the participants' actual long-term ability to control an MI-BCI. For the first time, performance predictors were not determined in a context of pure motor-imagery, since participants were asked to perform one motor imagery task -left-hand movement imagination- as well as two non-motor MI-tasks -mental rotation and mental subtraction-.

Four major results were obtained. The first is the strong correlation between MI-BCI performance and mental rotation scores. The second result is the definition of MODEL ♯1 which explained more than 96% of the variance of participants' MI-BCI performance. This model was composed of six factors: mental rotation, self-reliance, visuo-spatial memory span, tension, apprehension and the "visual/verbal" dimension of the learning style. The main flaw of MODEL ♯1 was its instability, revealed by the cross validation process. This instability could be due to the strong correlation of the mental rotation scores with MI-BCI performances, which could prevent other important factors from being expressed in the regression. Thus, the third major result is MODEL ♯2, from which the mental rotation factor was excluded. MODEL ♯2 explained more than 80% of MI-BCI performance variance and was composed of four factors: tension, abstractness, self-reliance and the “active/reflective" dimension of the learning style. This model appeared to be both stable and reliable to predict MI-BCI performance. Finally, the fourth result is the complementarity between MODEL ♯2 and mental rotation scores. Indeed, the only participants for whom MODEL ♯2 failed, by overrating their performances, were the participants with a very low mental rotation score.

Mental rotation scores reflect Spatial Abilities (SA) [156], i.e., the capacity to understand and remember spatial relations between objects. Mental rotation, and thus SA, are closely related with the three MI tasks considered in this study. First, it is obviously related with the mental rotation task. Second, [157] showed that children confronted with difficulties to perform arithmetics also had low spatial abilities. Third, the mental rotation test is actually used to evaluate motor imagery abilities in healthy subjects and patients with brain injuries [158]. The *self-reliance* and *tension* factors were probably selected due to their relationship with the nature of MI-BCI training which is a *distant learning*, i.e., a learning occurring in a context free of social interaction (the learner interacts with a computer, there are neither teachers nor students). Indeed, on the one hand, [159] showed that learners easily feel confusion, frustration and anxiety when confronted to distant education due to the lack of feedback from an instructor, compared to classic classroom education situations. Therefore, it seems relevant that learners with highly tense personalities encounter difficulties in distance learning tasks such as BCI training. On the other hand, in [160], autonomy, i.e., *self-reliance*, is presented as being of utmost importance in independent learning, and thus in distance learning. To summarize, it seems users with high “Tension" and low “Self-Reliance" traits may need a social presence and emotional feedback to improve their control performance. Finally, *Abstractness* refers to creativity and imagination abilities. It has been reported that creative people frequently use mental imagery for scientific and artistic productions [161] which could explain why participants with high abstractness abilities are more used to performing mental imagery.

2.3.2 Impact of the cognitive profile on pure Motor Imagery based BCI performances

The results of the previous study, and in particular the strong impact of SA on BCI performances, encouraged us to study whether such predictor was a generic predictor of performances, i.e., if it could also predict BCI performances on other experiments and tasks. We therefore investigated whether SA also had an impact on the performances obtained in a pure motor imagery BCI.

Protocol

20 BCI-naive participants (10 females; aged 24.7 $\pm$ 4.0 year-old) took part in this second study. Each participant had to learn to do 2 MI-tasks, namely imagining left- and right-hand movements, so that they were recognized by the system. Participants first had to complete a "calibration" run which aimed at providing the system with examples of EEG patterns associated with each of the MI-tasks. CSP spatial filters and an LDA classifier were calibrated on the data from this first run (see [137] for details) before being using to provide online feedback in the subsequent runs. Then, user training lasted 4 runs, each of them being composed of 20 trials per task. These trials were the same as those used in the Graz BCI protocol, already described in Section 2.2.1. Added to these MI-tasks, participants were asked to complete a mental rotation questionnaire, to measure their SA [152]. We then looked for correlations between the mental rotation scores obtained, reflecting spatial abilities, and the obtained BCI performances for this pure motor imagery-based BCI.

Results

In our analysis, we considered two different measures of MI-BCI performance: (1) the **peak** classification accuracy (measured at the time window of the feedback period for which the classification accuracy over all trials is maximal), which is the typical performance measure used with standard BCI training protocols such as the Graz protocol, see, e.g., [162] and (2) the **mean** classification accuracy over the whole feedback period of all trials. While Mental Rotation scores were not correlated with mean MI-BCI performance [$r=0.266$; $p=0.257$], they were correlated with the peak MI-BCI performance [$r=0.464$; $p=0.039$]. These results confirm the important impact of SA on mental imagery-BCI performance which was demonstrated in the previous section (see [120, 163]). More specifically, the positive correlation indicates that people with better spatial abilities (i.e., higher mental rotation scores in this instance) obtain higher MI-BCI control performance.

2.3.3 An overview of Psychological and Cognitive Factors impacting BCI performances

The two studies described above thus revealed some important psychological and cognitive factors impacting BCI performances. They reflect only part of the factors influencing performances though. Therefore we conducted a survey of the literature to review the latest developments in our understanding of the psychological and cognitive factors reported to influence MI-BCI performance (i.e., control accuracy). These factors can be divided into three groups. The first group includes the factors associated with the States of the user. Users' states are described by [164] as “temporary, brief, and caused by external circumstances”. The second group gathers the factors related to the users’ Traits, characterized as “stable, long-lasting, and internally caused” with respect to one’s environment and experience [164]. Finally, the third group comprises the factors that can be qualified neither as Traits nor as States, i.e., demographic characteristics, habits and environment-related factors.

Emotional and Cognitive States that Impact MI-BCI Performance

Some aspects of users’ states, and more specifically of their cognitive and emotional states, have been reported to influence their MI-BCI performance in terms of control accuracy. First, [165] have shown that mood (measured using a subscale of the German Inventory to assess Quality of Life - [166] -) correlates with BCI performance. On the other hand, both attention [167, 168, 169], assessed for instance by means of digit spans or block taping spans [167], and motivation [139, 51, 165] levels have repeatedly been shown to positively correlate with performance, both in the context of Slow Cortical Potential (SCP) and SMR based BCI. Furthermore, in their study, [165] suggested that higher scores in mastery confidence, i.e., how confident the participant was that the training would be successful, were correlated to better SensoriMotor Rhythms (SMR) regulation abilities, whereas higher rates of fear of incompetence were correlated to lower SMR regulation abilities. This last point has also been suggested in [170] for stroke patients taking part in BCI-based rehabilitation. More generally speaking, fear of the BCI system has been shown to affect performance [171, 172, 173]. In the same vein, control beliefs [173], i.e., participants’ beliefs that their efforts to learn would result in a positive outcome, and self-efficacy [51], which can be defined as participants’ beliefs in their own abilities to manage future events, have been suggested to play a role in BCI performance, in an SMR and an SCP paradigm, respectively. Mastery confidence, control beliefs and self-efficacy can be classed as context-specific states, i.e., states triggered each time a person faces a specific situation.

Personality and Cognitive Traits that Influence MI-BCI performance

On the one hand, several aspects of the cognitive profile have been related to BCI control ability. Memory span and attentional abilities have been shown to correlate with the capacity to regulate SCP in patients with epilepsy [167]. [139] also showed that attention span played a role in one-session SMR-BCI control performance. As mentioned before, our study shown that active learners seem to perform better than reflective learners [120] in a context of MI-BCI control. We also shown that abstractness, i.e., imagination abilities correlated with classification accuracy in an MI-BCI experiment [120]. Furthermore, [139] have proposed a model for predicting SMR-BCI performance

- which includes visuo-motor coordination (assessed with the Two-Hand Coordination Test) and the degree of concentration (assessed with the Attitudes Towards Work) - that reaches significance. More recently, [143] tested this model in a 4 session experiment (one calibration and three training sessions) within a neurofeedback based SMR-BCI context (i.e., involving no machine learning). Their results showed that these parameters explained almost 20% of SMR-BCI performance in a linear regression. However, the first predictor, i.e., visual-motor coordination, failed significance. With this model, the average prediction error was less than 10%. Finally, kinesthetic imagination and visual-motor imagination scores have both been shown to be related to BCI performance by [174]. Finally, as mentioned before, we found strong correlations between mental rotation scores and mental-imagery based BCI performance [120, 137].

On the other hand, concerning personality traits, [171] have obtained a positive correlation between a Locus of control score related to dealing with technology and the accuracy of BCI control. As previously described, we shown that tension and self-reliance were related to MI-BCI performance in a model also including abstractness abilities and the active/reflective dimension of the learning style [120]. This model enabled prediction of more than 80% of the between-participant variance in terms of performance with an average prediction error of less than 3%.

Other Factors impacting MI-BCI Performance

Some other factors that have also been related to the ability to control a BCI, cannot be classified as either traits or states. These factors can be divided into three categories: (1) demographic characteristics, (2) experience/habits and (3) environment. Concerning the first point, demographic characteristics, age and gender have been related to SMR-BCI performance [141]: women being more capable than men and over 25 year-old being more competent than their younger counterparts. On the other hand, some habits or experiences have been shown to increase SMR-BCI control abilities [142, 141]. More specifically, playing a musical instrument, practicing a large number of sports, playing video games [141], as well as spending time typing and the ability to perform hand and arm or full-body movements [142] positively impact SMR-BCI performance. However, the consumption of affective drugs seems to have the opposite effect [142]. Finally, the user's environment, and more particularly the quality of caregiving for patients, has been suggested in an anonymous report to play a role in SMR-BCI performance [136].

Summary

To summarize, the predictors of MI-BCI performance can be gathered into the three following categories, as depicted in Figure 2.6:

STATES	EMOTIONAL STATE	♣	Mood (Nijboer et al. 2008)
	COGNITIVE STATE	♣	Attention level (Grosse-Wentrup et al., 2011; Grosse-Wentrup and Scholkopf, 2012)
		♣	Motivation (Hammer et al., 2012; Neumann and Birbaumer, 2003; Nijboer et al., 2008)
		♠	Mastery confidence (Nijboer et al., 2008)
		♠	Fear of the BCI (Burde and Blankertz, 2006; Nijboer et al., 2010, Witte et al., 2013)
		♠	Control beliefs (Witte et al., 2013)
		♠	Fear of incompetence (Kleih et al., 2013; Nijboer et al., 2008)
		♠	Self-efficacy (Neumann and Birbaumer, 2003)
TRAITS	PERSONALITY	♠	Locus of control for dealing with technology (Burde and Blankertz, 2006)
		♠	Tension (Jeunet et al., 2015)
		♠	Self-reliance (Jeunet et al., 2015)
	COGNITIVE PROFILE	♣	Attention span (Hammer et al., 2012)
		♣	Attentional abilities (Daum et al. 1993)
		♣	Attitude towards work (Hammer et al., 2012)
		♣	Memory span (Daum et al., 2013)
		♦	Visual-motor coordination (Hammer et al., 2014, 2012)
		♦	Learning style: active vs. reflective learners (Jeunet et al., 2015)
		♦	Kinaesthetic imagination score (Vuckovic and Osuagwu, 2013)
		♦	Visual motor imagination score (Vuckovic and Osuagwu, 2013)
		♦	Mental rotation scores (Jeunet et al., 2015, Jeunet et al., 2016)
		♦	Abstractness (Jeunet et al., 2015)
OTHER FACTORS	DEMOGRAPHIC DATA	·	Age (Randolph, 2012)
		·	Gender (Randolph, 2012)
	EXPERIENCE	♦	Playing a music instrument (Randolph, 2012)
		♦	Practicing sports (Randolph, 2012)
		♦	Playing video-games (Randolph, 2012)
		♦	Hand & arm movements (Randolph et al., 2010)
		♦	Time spent typing (Randolph et al., 2010)
		♦	Full body movements (Randolph et al., 2010)
		♣	Consumption of affective drugs (Randolph et al. 2010)
	ENVIRONMENT	·	Quality of caregiving (Kleih and Kübler, 2015)

Figure 2.6: Summary of the different predictors which have been related to MI-BCI performance in the literature, i.e. the predictors related to the user-technology relationship (orange spades), to attention (green clubs) and to spatial abilities (blue diamonds).

– Category 1 - The user-technology relationship & the notion of control (in orange - spades, see Figure 2.6): indeed, based on the literature, it appears that people who apprehend the use of technologies (and more specifically the use of BCIs) and who do not feel in control, experience

more trouble controlling BCIs.
- Category 2 - Attention (in green - clubs, see Figure 2.6): this category includes both attentional abilities (trait) and attention level (state). The latter can fluctuate with respect to different parameters such as environmental factors, mood or motivation. Both these aspects of attention have been repeatedly evoked as being predictors of BCI performance.
- Category 3 - Spatial Abilities (in blue - diamonds, see Figure 2.6): many predictors depicted in the previous brief review are related to motor abilities (e.g., 2-hand coordination, sports or music practice) or to the ability to produce mental images (e.g., kinaesthetic imagination scores). These predictors can be gathered under the label of "spatial abilities".

2.3.4 Towards adapted and adaptive MI-BCI training

The studies presented above have highlighted the huge impact of spatial abilities on MI-BCI performance. Our future work will consist in designing new kinds of MI-BCI training protocols aiming at improving users' spatial abilities (SA), prior to MI-BCI use [163]. Concretely, based on his/her basic spatial abilities, the user will be provided with specific SA-training exercises. We have very recently made a first step in that direction by designing and validating a set of such SA training exercises that can be integrated in an MI-BCI training scheme [175]. We are currently in the process of comparing this integrated SA training to standard BCI training to see if it can lead to increased performances.

Furthermore, in order to take into account the personality factors related to MI-BCI performance, a virtual learning companion will be developed. It will be able to provide the user with (1) cognitive support (e.g., by proposing examples) in the case of students with low abstractness abilities, (2) emotional and social support, notably social presence by giving advices and/or encouragements during the training procedure, for users with high "tension" and low "self-reliance" scores. We are currently designing and testing such a learning companion.

Such improved training protocol, that will be adapted to each user could potentially greatly increase the acceptability and accessibility of MI-BCI based technologies. Nevertheless, there is still a number of areas in which BCI user training can be improved independently of the user profile. In the following we present our works towards improving BCI feedback and training tasks.

2.4 Improving BCI training tasks and feedback

Given the identified limitations of current BCI user training approaches, we have also conducted works to improve such training approaches. In particular, we have design new feedback types and new feedback environments, designed to satisfy guidelines and principles from educational sciences and human learning theories. Such principle indeed recommend to provide a feedback that is multimodal [131, 133], explanatory [130, 129], motivating [129, 128], as well as exploiting social interactions [134, 135, 136]. We therefore present below our work in these directions. In particular we present a motivating feedback in a social context with a multi-user BCI training in Section 2.4.1, a continuous tactile feedback for BCI in Section 2.4.2, and first steps towards an explanatory feedback in Section 2.4.3.

2.4.1 Multi-subjects BCI training

One of the limitations of standard BCI training training and feedback is that it is not motivating and does not benefit from a social context or social interactions, which are also essential for successful learning and good task performances [134, 135, 136]. To address these limitations, we have therefore design and studied a multi-user BCI game - called BrainArena - to train two users to control a BCI by working in cooperation in a motor imagery-based game [176]. We describe the game below and the evaluation performed to compare single user BCI control and multi-user BCI control.

BrainArena - a multi-user BCI game

BrainArena is a multi-player football game controlled by hand motor imagery. The objective of the two users is to imagine left or right hand movements to move a virtual ball towards a goal located on the left or right side of the screen, respectively. Each user has his/her own EEG acquisition and tuned BCI pipeline. During the game, a ball was displayed at the center of the screen on a black background. Goals were symbolized by two triangles on each side of the screen (see Figure 2.7). A green cross was displayed in the center, with green or blue feedback gauges extending left or right during the game sessions.

The feedback gave two complementary pieces of information. First, the real-time feedback on the “intensity” of the commands given by the users was provided: a gauge which went left or right depending on whether left or right motor imagery was recognized, and whose length represented the actual intensity of the command (i.e., the distance of the feature vector to the LDA hyperplane). When two users were playing together, 3 gauges were displayed: the two single-user feedbacks plus a multi-user gauge in the middle for the resulting overall command, summing both users’ commands. The lengths of the user gauges were directly proportional to the normalized output of the LDA classifier(s) used in the BCI process. The second form of feedback was that the ball could roll horizontally when pushed by the mental commands (i.e. left or right hand motor imagery), acting like a cumulative feedback for the user.

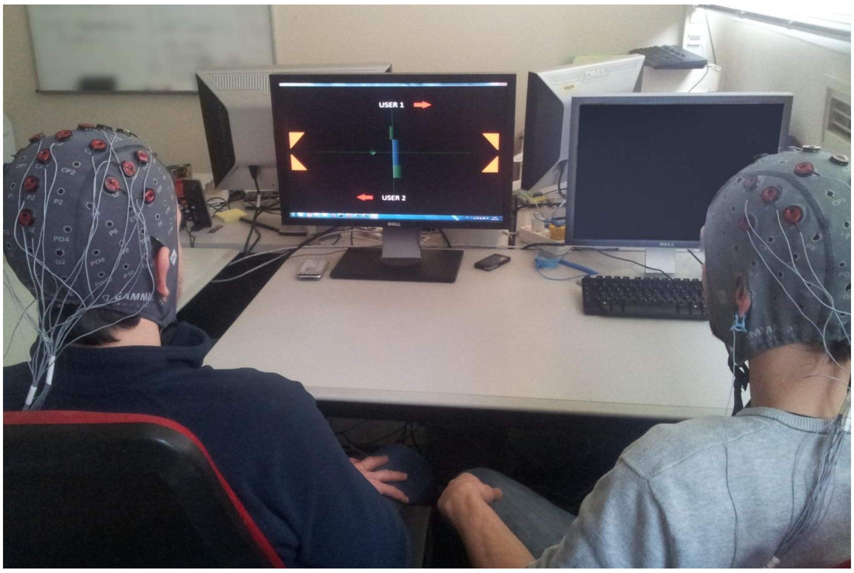


Figure 2.7: Two users playing to the BrainArena Motor Imagery BCI game.

The real-time gauges could be viewed as a representation of these push forces applied to the ball, each push being added to the previous ones to move the ball. This ball had a physics-based behaviour, thus when pushed it acquired a velocity. Moving the ball to a given goal was the objective of the game, and would increase the users scores.

We designed three different paradigms, one single-user and two multi-user interactions. The Solo mode involves only one user, thus one BCI pipeline. The system asks the user to score a goal on the left or right side of the screen, by performing an imaginary movement of the corresponding hand. In multi-player gaming, people are expected to work together to achieve a goal and/or work against each other to be the best. Thus we designed two versions of our multi-user BCI game: a collaborative mode, where players are supposed to join forces to achieve the goal and improve their score, and a competitive version where they must perform better than their opponent. Here we describe only the collaborative version which was studied after. The Collaborative version received inputs from two BCI systems. The two users shared the same objective: moving the ball to the left or right goal. The application displayed the feedback gauge of both users. Between users' individual feedbacks, the multi-user feedback was presented in blue as the sum of both feedbacks (see Figure 2.7). The ball was pushed by this multi-user feedback. In terms of signal processing, the BCI used were classically designed around the CSP and LDA algorithms (see [176] for details).

Evaluation

We aimed to study the impact of a collaborative multi-user situation on the performances of two users connected through BCI to the same video game. We therefore compared a single player version (SOLO) to a collaborative version (COLLAB) of the game. The classification accuracy during the online sessions was used as a performance metric. The reported performances are the maximal classification accuracy over the trial duration, as done by the Graz group [92, 177]. The population consisted of 20 volunteers (15 males and 5 females), all of them naive users of BCI technologies. From this group 10 pairs were formed.

The experiment consisted of 5 runs, with 40 trials per run (20 for each MI task). The first run was the acquisition of a training set for the CSP filters and LDA classifier. During this first session no feedback was displayed. The following sessions were either in the SOLO or COLLAB condition (2 SOLO, 2 COLLAB). Preliminary testers reported a better understanding of the instruction if they started in the SOLO condition. Therefore, the first session was always in the SOLO condition, the condition order for the other sessions being randomly chosen.

Results

The mean accuracy for the SOLO condition was 71.3% while it was 73.9% for the COLLAB condition. This difference was not found to be significant with a paired t-test although it did show a trend ($p = 0.06$). There was no observed learning effect over the 4 sessions. We divided the subject pool into 2 subgroups, according to their performance levels. The *Winner* subgroup consisted of the dominant participants of each pair (best mean overall accuracy). The *Loser* subgroup was the other half, with the worst mean accuracy of each pair. In the *Winner* group the mean classification accuracy in SOLO condition was 75.0%, and 80.0% in the COLLAB condition. This difference was found to be significant ($p \leq 0.01$). The *Loser* group showed no significant differences between the two conditions (SOLO: 67.5%, COLLAB: 67.8%).

Discussion

Our evaluation on 20 naive subjects compared the single-user situation with the multi-user situation using the collaborative condition. Although the mean classification performance was not significantly better in collaborative condition, it showed a tendency (p = 0.06), which will have to be confirmed in further studies. However, when analyzing separately the best performing users and the worst performing ones from each pair, we found a significant difference between collaborative and single-user for the best performing user only. This means that operating a BCI in a multi-user context is possible without any performance drop, and may even increase the classification performances of the best performing users. This confirms that exploiting social interaction (here collaboration) and/or motivating training environment (here a 2-player BCI game) can indeed enhance BCI performances and make the training more enjoyable for our BCI users, as theoretically suggested [134, 135, 136, 130]. The proposed multiplayer BCI game is one possible approach to do so. In the future, it would be worth further exploring those ideas to design training protocol that are motivating and in a social context, and that can improve BCI performances of all users, not only of the best of each pair.

2.4.2 Continuous tactile feedback

As mentioned before, most MI-BCI studies to date involved visual feedback to inform the user about the MI task recognized by the system. Yet, this visual feedback is difficult to assimilate when integrated with the visual layout of the primary interactive application that it supports [178]. Indeed, the visual channel is often overtaxed in interactive environments [179]. Thus, integrating the visual feedback into the application increases the number of visual search tasks. This is a typical branching condition [180] where users have poor performance in searching for visual object [181]. On the other hand, tactile feedback, although popular in other areas of HCI, has not received much attention for MI-BCI despite its advantages such as: (a) freeing the visual channel in order to reduce cognitive workload [179], (b) maintaining a certain amount of privacy, as it is more difficult to be perceived by the surroundings than the visual or auditory ones, and (c) the possibility to be used in a wide range of interactive tasks, such as in gaming conditions. Using tactile feedback will separate the application channel (visual) from the MI-BCI feedback channel (tactile), thus potentially improving the branching condition of the application. Finally, as mentioned earlier in this chapter, exploiting multimodality can increase learning efficiency [131, 133]. This should consequently increase the user's performance and system's efficiency.

We therefore explored the benefits of providing a continuously updated tactile feedback to improve MI-BCI users' performance in an environment containing visual distracters (see Figure 2.8) [182]. Indeed, BCIs are inherently developed to promote interaction. Yet, most MI-BCI studies test their feedback efficiency (1) in a laboratory context, i.e., with no distracters and (2) with no side task, while in real applications such as games users would have to perform multitasking. Thus, the efficiency of these feedbacks cannot be guaranteed in an interaction and multitasking context. This is why we study our tactile feedback's efficiency by comparing it to an equivalent visual feedback, (1) in a context including visual distracters (to mimic an interaction environment) and (2) by adding a counting task (to evaluate the cognitive resources needed to process each kind of feedback). Our tactile system is in the form of a wearable glove that integrates five vibrotactile actuators for each hand to provide continuous tactile feedback to the user regarding the classifier output. This expands the user's feedback bandwidth while reducing the visual cognitive load. We describe below the design of the tactile and visual feedbacks, and then the experiment conducted to compare them, before discussing the study results.

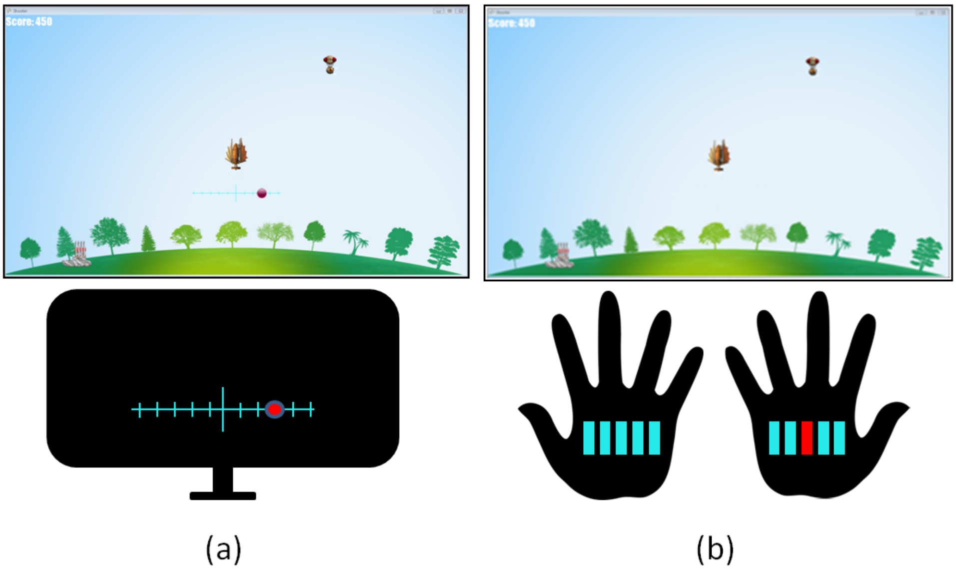


Figure 2.8: Illustration of the combination of an appealing training environment and a vibrotactile feedback provided on the palm of the hand using an array of vibrators (right, compared to a standard visual feedback, left)

Visual and Tactile Feedback design

Temporally Continuous Tactile Feedback Our goal was to represent the BCI feedback, i.e. the classifier output, via the tactile channel as closely as possible to the standard visual feedback (in which the output is represented as a bar varying in length and direction - see Section 2.2.1). The MI-BCI system relies on left- and right-hand MI. Thus, we decided to give tactile feedback to the hands to maintain the control-display mapping [183] between the intended user actions (MI) and the sensory information perceived by the user (the tactile feedback). Indeed, control-display mapping has been show to be necessary so that tactile feedback is efficient [183]. The large surface of the palm (the average width is 74 mm for women, 84 mm for men) makes it possible to create a tactile display suitable for representing the MI-BCI classifier output (see Figure 2.8). Indeed, considering the two-point threshold of the palm (about 8 mm [184]), the width of the actuators, 8 mm, and the fact that we wanted our design to be suitable for most of the users (and thus narrower than the average palm width, 74 mm), we determined that we could put 5 motors maximum on each hand. Thus, we divided the classifier output range (here $[-0.5, 0.5]$ with an SVM) into 10 discrete levels, with 5 levels on the left and 5 levels on the right hand. Vibrations on the left/right palm corresponded to the recognition of a left/right hand MI by the classifier, respectively. With the palms being facing upwards, vibrations near the thumbs corresponded to high confidence levels (close to $|0.5|$) while vibrations near the little finger corresponded to low confidence levels (close to 0). For details on the hardware used to build such tactile gloves see [182]. We also conducted a user study to identify the best intensity and pattern of activation of the motors in terms of user experience, see [182] for details.

Visual Feedback Standard visual feedback corresponds to a continuous bar varying in length and direction. To make both the visual and tactile feedback as similar as possible, and because the tactile feedback has been spatially discretized (classifier output range of $[-0.5, 0.5]$ divided into 10 discrete levels), we also discretized the standard bar in the same way. Thus, the feedback was displayed as a red cursor on a cross, with 5 ticks on the left and 5 ticks on the right side (see Figure 2.8). The cursor was on the left/right side of the cross when a left/right hand MI was recognized, respectively. Moreover, the cursor moving to the extremities of the cross represented high confidence values.

Evaluation

In order to include the distracters and the counting task to the MI-BCI task in a consistent environment, we modified the standard MI-BCI training protocol described in Section 2.2.1. The standard arrows pointing left or right to inform the user he has to perform a left or right-hand MI have been replaced by a spacecraft the goal of which was to protect its planet by destroying bombs coming from the left or right (controlled by performing left- or right-hand MI, respectively) (Figure 2.8). Besides, the distracters were appearing randomly in the form of (1) a missile, which was launched in a vertical direction from a tank, (2) a rabbit crossing from the left to the right, or (3) a cloud crossing from the right to the left (Figure 4). Each distracter appeared for a similar amount of time (approximately 2.5s).

Eighteen healthy volunteers (5 women; aged 27.6 ± 4.8) participated in the study. Some of them had previously experienced vibrotactile feedback. However, none of them had previous experience with MI-BCI. Nine participants were provided with visual feedback, and the other nine participants were provided with vibrotactile feedback during the whole experiment.The experiment was divided into 6 runs, each of 7 minutes duration. The first run was used to train the MI-BCI classifier. The remaining 5 runs were used for the user training and data recording. Each run was composed of 40 trials: 20 left-hand MI and 20 right-hand MI trials, randomly distributed.

During the experiment, the participants had to control a spacecraft (shown at the center of the screen in Figure 4) by performing left- or right-hand MI tasks to make it move left or right, respectively. The goal of this spacecraft was to protect the planet against bombs falling down on the planet. Thus, when a bomb was falling off the left/right side of the screen, participants had to perform left/right-hand MI in order to make the spacecraft move left/right, face the bomb and destroy it. Each trial was lasting around 8s and had the same structure, described hereafter. At the beginning, the spacecraft was in the middle of the screen for 3s. Then the instruction was given to the participant as a bomb appearing either at the top left or right of the screen and moving vertically towards the planet. This instruction informed the participant about the command to perform: a right-hand or a left-hand MI, in order to move the spacecraft to the right or to the left, respectively, face and destroy the enemy. 1.25s after the appearance of the bomb, the MI-BCI classifier output was provided to the participant continuously for a duration of 3.75 seconds, either in the form of a moving cursor on a visual cross at the lower center of the screen, or as vibrotactile feedback at the palm. At the end of the feedback period, the mean classifier output was calculated and the spacecraft was moving to the left or right, depending on whether this value was negative or positive, to intercept (or not) the bomb. Furthermore, as explained before, during each trial, one or more distracters were appearing between the moment when the enemy was displayed and the moment when the spacecraft started to move in order to catch the bomb. Each distracter type appeared at most once during each trial. In each Run, which consisted of 40 trials, each distracter type appeared at least 15 times and at most 25 times. At the beginning of each Run the participants were asked to count how many distracters of a specified type appeared, and to report this number at the end of the Run.

In terms of signal processing, the BCI was built around a classical CSP spatial filtering and an SVM as classifier. The SVM was providing an output between -0.5 (left hand MI) to 0.5 (right hand MI), see [182] for details. At the end of the trial, the score was updated according to the formula:

$$\begin{aligned} NEW\ SCORE =& CURRENT\ SCORE \\ &+ CLASS\ LABEL \times CLASSIFIER\ OUTPUT \times 200 \end{aligned}$$

The CLASS-LABEL was -1 if a left-hand MI was recognized and $+1$ if a right-hand MI was recognized. The CLASSIFIER-OUTPUT was the mean classifier output value calculated at the end of the trial. The MI score corresponded to the sum of the scores obtained in each trial. Furthermore, at the end of each run, the participant was asked to report the number of distracters (rabbits, clouds or rockets) he counted. If this number was correct, the participant was rewarded with 200 points being added to the MI score. If the error was in the order of ± 1, the score remained unchanged. Otherwise, 200 points were subtracted from the MI score. The final score corresponded to the sum of the MI scores for the 40 trials of the run and the counting task score. While arbitrary, this metric enabled to consider and give a significant weight to both the MI score and the counting task which allowed to evaluate the feedback relevance for both these aspects.

Results

The main measurements of interest are (1) the final score (the sum of the MI task and the counting task scores), (2) the MI score alone, and (3) the absolute value of the difference between the counted and the actual number of distracters. These measures were analyzed using three two-factor (independent) ANOVAs. We performed a 2-way ANOVA so that we can analyze the interaction between both variables. Analyzes have been performed on 17 participants: 8 in the visual condition and 9 in the tactile condition. The data from one outlier participant have been removed as his final score (1628.8 ± 630.5) differed considerably from his group mean final score (183.0 ± 559.5).

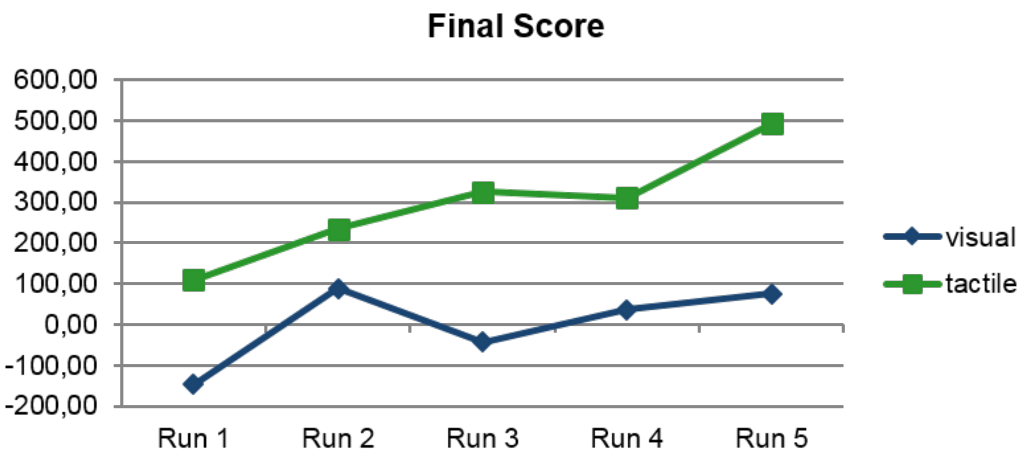


Figure 2.9: Average of the final scores: sum of the MI task score and the distracter counting task score (reward and penalty).

The two-factor ANOVA on the final score shows a main effect of the Feedback-Condition (visual vs. tactile) $[F(1, 15) = 6.327, p < 0.05]$, a main effect of the Run $[F(1, 15) = 3.961, p < 0.01]$ but no Run × Feedback-Condition interaction $[F(1, 15) = 1.476, p = 0.243]$. The Feedback Condition effect is due to participants in the tactile feedback group having significantly better results than participants in the visual feedback group. Furthermore, concerning the Run main effect, post-hoc analysis shows a significant increase of performance between Run 1 and Run 5 ($p < 0.005$) (Figure 2.9) which reveals the learning effect of the performed motor-imagery task, as indicated by the large effect

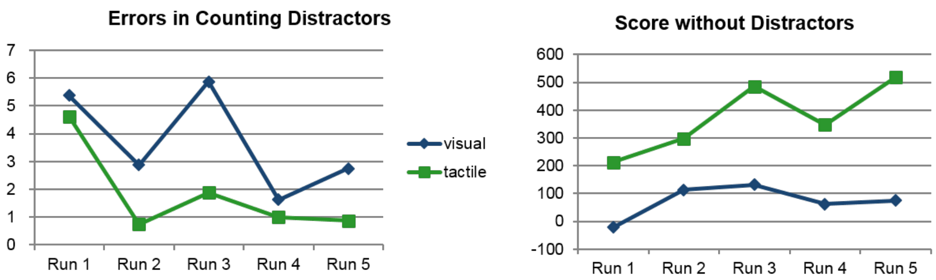


Figure 2.10: Left: Average of the MI scores without the counting task (reward and penalty). Right: Average of the distracter errors (difference between the counted and the actual number) for the counting task as a function of Run number and Feedback Condition.

size. The two-factor ANOVA on MI scores (Figure 2.10, left) shows strong tendencies towards a Run main effect $[F(1,15) = 3.961, p = 0.065]$ and towards a Feedback Condition effect $[F(1,15) = 4.063, p = 0.062]$, as well as no interaction between these two factors $[F(1,15) = 1.207, p = 0.289]$. These results indicate a strong tendency towards a better MI score with tactile feedback than with visual feedback and a tendency towards an improved MI score across the Runs. The two-factor ANOVA on the counting task (Figure 2.10, right) shows a main effect of the Run $[F(1,15) = 9.806, p < 0.01]$ but no main effect of the Feedback Condition $[F(1,15) = 2.860, p = 0.111]$ and no Run $\times$ Condition interaction $[F(1,15) = 0.000, p = 0.990]$. Thus, the participants improved their performance across the Runs for the counting task. Indeed, post-hoc analysis shows a significant difference between Run 1 and Run 4 ($p < 0.001$) and Run 1 and Run 5 ($p < 0.005$) performances.

Conclusion

Our study suggested that it was possible to provide MI-BCI users with a relevant continuous vibrotactile feedback while they are performing MI tasks, and that this tactile feedback can improve BCI control reliability in a multitasking context (as compared to an equivalent visual feedback) [182]. It suggests that providing feedback through another modality than the visual one, but with the same content has advantages: it tends to improve users' BCI control, frees the visual channel and thus some cognitive resources to perform other tasks, as it was suggested by human learning theories [129, 128]. Besides, receiving a continuous tactile feedback consistent with the motor imagery tasks being performed is probably more natural and intuitive than a visual feedback. In this study, only the feedback form and modality was investigated though. Yet, much work has to be done on the feedback content so that it can be really relevant. As mentioned before, among others the feedback content should be explanatory, supportive and meaningful. In the next section, we present our first steps towards designing such an explanatory feedback.

2.4.3 Towards explanatory feedback

Theoretically (see Section 2.2.2) feedback should be explanatory, motivating and meaningful, whereas current BCI feedback is usually boring, corrective only and difficult to understand. In this study, we explored different EEG signal features to be used as a richer, explanatory BCI feedback [121]. In order to determine which additional information could be presented as feedback we analyzed data from a previous experiment (presented in Section 2.3.1) to find EEG features that are indicative of performance and more easily understandable for the users.

Identifying Feedback Features

We explored different EEG features that might be presented as feedback to support BCI user training. These features were assessed according to their ability to predict performance by using data from the experiment described in Section 2.3.1, during which we trained our users to perform left hand motor imagery, mental subtraction and mental rotation of a geometric figure.

The additional feedback information should follow three main objectives. Firstly, it should be available online during the experiment and it should be possible to update it regularly. Secondly, the feature should have a clear connection to BCI performance so that information about it might help people to improve performance. Also, the direction of the connection should be clear, i.e. the user should know if they have to up– or down–regulate the respective feature to improve performance. Thirdly, the feature should not involve any complex computations that users might not understand and/or have an intuitive explanation. With these three objectives in mind, we notably determined the following two features:

Muscle bandpower: One important aspect for the EEG signal recording in general and for the use of a BCI in particular is the reduction of muscular artifacts, as it can produce electromyogram (EMG) activity that interferes with the EEG signals [185]. To measure muscular tension we used three electrodes from the front (F3, Fz and F4) and three electrodes from the back of the head (PO7, Oz and PO8) and determined the mean bandpower in a frequency band from 40 to 70 Hz which is associated with muscular activity. The hypothesis for this feature was that it should be negatively correlated to BCI performance, i.e. the lower the bandpower in the muscle frequency band the better the performance in the BCI task.

Gamma bandpower: Another important aspect influencing performance during BCI experiments is the level of attention the user is paying to the task, as presented before (see Section 2.3.3). As in [186] we used a linearly–constrained–minimum–variance (LCMV) beamformer to determine the relevant brain region in which γ power - reflecting high level attention [168, 169] - is associated with BCI performance. A beamformer can be seen as a spatial filter which in this case was aimed at the superior parietal cortex, a region that was found to best predict BCI performance in [169]. In the present study a 5s long baseline period from the beginning of the calibration run was used to compute the beamformer for every subject individually. The resulting spatial filter was then applied to the signals from the four feedback runs before the bandpower in the γ frequency range (55-85 Hz) was calculated.

To evaluate these features regarding their relation to BCI performance two selection criteria were tested. First, we looked for correlations between the respective feature and BCI performance (mean classification accuracy over the feedback period) across subjects. However, in order for the feedback to be useful online during the experiment, we also looked for correlations with performance within-subjects, on smaller time windows (i.e., 1s windows as during online processing). To take into account not only the class label predicted by the classifier, but also the certainty of the classifier in its prediction, we followed the approach in [168]. They proposed a score - which we denote here as LDA-score - which is calculated by taking the absolute classifier output (output of the LDA for the winning class in our one-versus-the-rest multiclass LDA scheme) in case of a correct time window, and by multiplying the absolute classifier output by -1 in the case of a wrongly classified time window. By this means a continuous–valued performance measure on the basis of single time windows is obtained which can then be correlated with the values of the respective feedback feature.

Results for feedback features: For both features we could observe a significant negative correlation between the feature and BCI performance, with $\rho = -0.3391$ ($p = 0.008$) for the muscle bandpower and $\rho = -0.4243$ ($p = 0.0007$) for the gamma bandpower. Correlations between the two features and the LDA-score for each subject are displayed in Figure 2.11. A star indicates that the correlation is significant ($p < 0.05$, t-test). For the γ bandpower analyses the direction of the correlation is almost consistent across subjects, i.e., the feature is almost always negatively correlated to the performance score. For the muscle bandpower the direction of the correlation is more subject–specific. While it is negative for most subjects, there are three subjects who show a significant positive correlation. In all cases the strength of the correlation varies across subjects.

Both features explored here were significantly correlated with BCI performance across subjects and runs. Thus, they can – to a certain extent – explain

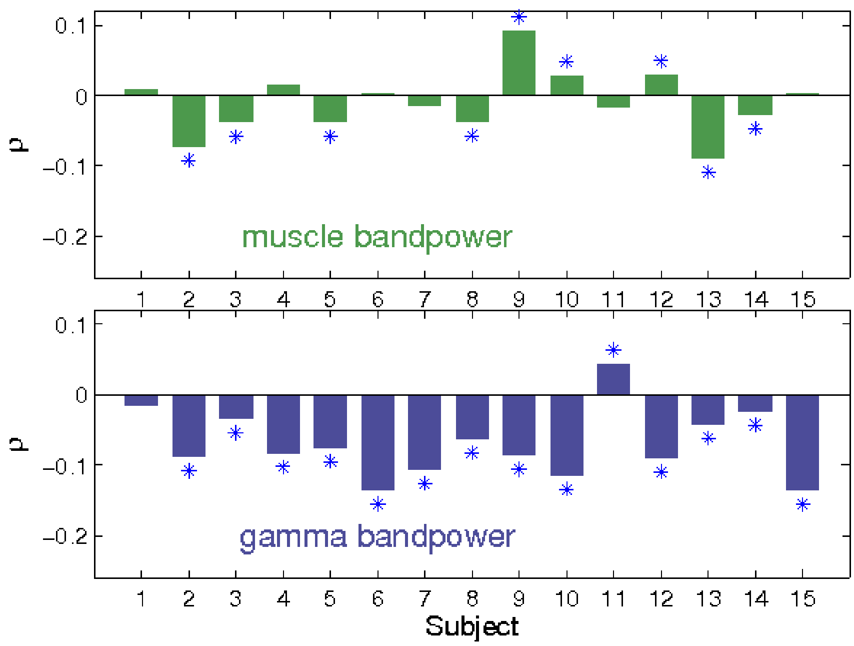


Figure 2.11: Correlation between the LDA-score and the tested feedback features per subject. Stars indicate significant correlation coefficients ($p < 0.05$).

the variability in BCI performance between different runs and subjects. In the case of the muscle bandpower the correlation was negative which means that it went in the expected direction, i.e. the more tensed a subject the lower the BCI performance. The analyses of γ bandpower also yielded a negative correlation which is not what we expected. Actually, the correlation between the muscle bandpower and the gamma bandpower was $\rho = 0.57$ ($p < 0.00001$), which suggests that our gamma bandpower feature was probably measuring a mix of muscle activity, and, hopefully, attention.

Based on the results from the feature evaluation we decided to test the muscle bandpower to use it as feedback in an actual BCI experiment. Indeed, the muscle bandpower being linked to performance, informing subjects about their level of muscular tension might help them to improve BCI performance. Moreover, users can easily interpret the muscle band-power as a measure of muscular tension and muscular artifacts, and thus regulate it. The next section reports on the evaluation of an online BCI using the muscle bandpower as additional feedback feature to help users acquire BCI control.

BCI Experiment with additional feedback

The experiment structure was similar to the one described in Section 2.3.1, the data from which were used to assess the feedback features. It also consisted of four feedback runs preceded by a calibration run with sham feedback, with the same trial timing. The only difference lies in the administration of a new feedback that was provided before the start of the visual bar feedback. It was presented from 1s after the appearance of the fixation cross for a duration of 3s, i.e. it ended right before the blue bar feedback began. The new feedback was given in auditory form as a sound that was played from loudspeakers. In a pilot experiment different sounds were tested. Most subjects preferred a relaxing sound which was therefore chosen. It was presented during the 3s long period whenever the muscle bandpower exceeded a certain threshold. Participants were asked to relax their muscles and to avoid any movement in order to make the sound stop during the 3s period. They were informed that the sound stopped right before the visual feedback started in any case.

The auditory feedback was only given in the three last feedback runs. The first feedback run was exactly the same as described in Section 2.3.1. Data from this run was used to determine the threshold which was computed as the 60^{th} percentile of the distribution of the muscle bandpower during this first feedback run. Thresholds were subject–specific. A pilot experiment showed that the distributions of the muscle bandpower for a separate baseline period as well as for the calibration run were too different from the bandpower distribution during the feedback runs. Therefore, the first feedback run was used to determine a realistic threshold.

10 BCI–naive participants took part in this experiment (6 males, 4 females, mean age 29.1, range 23–42). The data from ten subjects was selected from the previous experiment (see Section 2.3.1) so that it matched in terms of performance during the first feedback run. They constituted the control condition.

Results

Figure 2.12 shows the mean performance over subjects for the three runs and compares it to the mean performance over 10 control subjects from the previous experiment. Note that the first run was the same for both conditions while run two to four included the additional auditory feedback in our experiment. Control subjects were chosen such that the average performance in the first feedback run (59.89%), was as similar as possible to the performance in the first run of our experiment (59.95%, p=0.9830).

Although the performance is slightly higher with additional auditory feedback in run 3, a two–way analysis of variance (ANOVA) for repeated measures with the factors *Condition* (C_2: auditory vs no auditory feedback) and *Runs* (R_3: run 2, 3 and 4) did not reveal any significant main effect of the *Condition* [F(1,18)=0.369, p=0.551], the *Run* [F(1,18)=0.480, p=0.497] nor a *Condition*Run* interaction [F(1,18)=0.003, p=0.960]. To evaluate the effect of the additional feedback on muscular relaxation, 10 subjects were chosen from the previous experiment such that they matched in terms of muscle

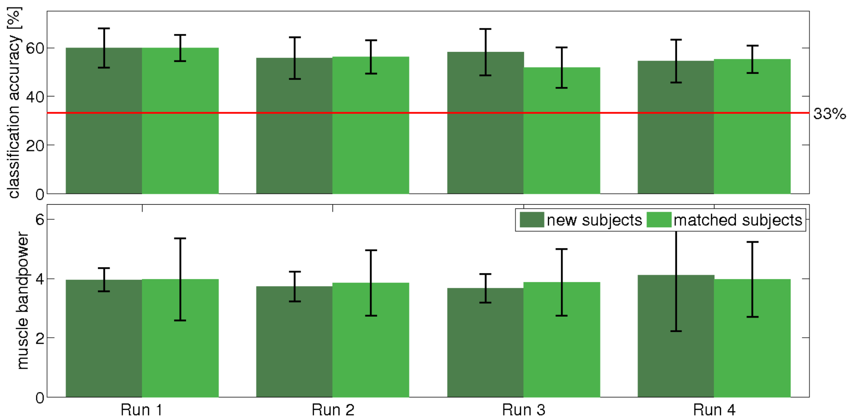


Figure 2.12: Top: mean classification accuracies for subjects from the experiment and the control group. Subjects from our experiment received an auditory feedback about muscular tension in addition to the visual moving bar feedback. The control group only received the visual moving bar feedback. The red line indicates chance performance for this three–class experiment. Bottom: corresponding bandpower in the 40–70 Hz frequency band. In this case control group is chosen to match the muscle bandpower during Run 1.

bandpower during the first run (p=0.9855). In line with the slightly better performance the muscle bandpower is a little lower for subjects with additional auditory feedback in run 3 (see Figure 2.12). Again a two–way ANOVA for repeated measures showed no main effect of the *Condition* [F(1,18)=0.017, p=0.897], the *Run* [F(1,18)=0.845, p=0.370] nor a *Condition*Run* interaction [F(1,18)=0.220, p=0.644].

Discussion and conclusion

We did not find any significant effects of the additional auditory feedback, neither in terms of performance nor in terms of muscle bandpower. It may be due to its duration (3s), which might have been too short for the user to understand the feedback and react to it by relaxing. This duration was chosen to fit the trial timing from the previous experiment, used as a control condition. It would thus be interesting (1) to study the effect of a longer auditory feedback period and (2) to postpone the start of the BCI task until the feature has reached a desired value as suggested in [169, 187]. In our case, this would mean to wait with the presentation of the BCI task until the user is relaxed enough, i.e. the muscle bandpower is below a predefined threshold. Besides, BCIs being co–adaptive systems both the user and the machine might be responsible for low performance. When evaluating the different feedback features according to their ability to predict performance we implicitly assumed that the user was responsible for incorrectly classified time windows, e.g. he was not relaxed enough or was not paying enough attention to the task. Yet, it could also be due to the computer, e.g., to an imperfect signal processing or classification algorithm. In the future, distinguishing between human and computer errors should help to identify more specific feedback features.

Altogether, this study aimed at exploring various EEG features that could be used as explanatory feedback for BCI training. The challenge was to find a feature that could be used online. As such we explored a measure of muscle tension from the frontal and neck areas and the γ bandpower as a measure of the attentional level. Both appeared to be correlated to BCI performance. In the case of the γ bandpower it might be problematic to distinguish γ power associated with attention from muscular artifacts. However, if a contamination by EMG activity can be excluded it might be a useful feature for BCI feedback.

Furthermore, we tested the muscle tension feature in a BCI experiment to investigate whether it is useful to improve user performance. Although no improvement was found in the experiment using feedback about muscular relaxation, adjustments to the experimental design might help to make this feature useful to improve BCI performance, as discussed above. Moreover, we could show that it is possible to use an additional feedback during a BCI experiment without deteriorating performances. We hope this study could be a first step towards designing explanatory feedback for BCI with the objective of improving BCI training and thus BCI reliability.

2.5 Discussion and perspectives on BCI user training

To summarize our contributions on BCI user training, we have first shown that current standard BCI user training approaches were inappropriate as they did not satisfy guidelines and principles from educational science and human learning theories, and as even in practice, they prevented several users from learning simple motor tasks. With the longer term objective to design adapted BCI training, which is theoretically recommended, we looked for psychological predictors of BCI performance, to determine how the user's profile impacted BCI performances. Our studies notably revealed the crucial role of Spatial Abilities (SA) in BCI control performances, both for mental imagery-based BCIs and pure motor imagery based-BCIs. In addition to SA, a survey of the literature enabled us to identify that attention and relationship with technology are two other major predictors of performances. Still to improve BCI user training, we have explored new feedback types, in particular continuous vibrotactile feedback and multi-user feedback, both leading to increased BCI performances. We have also made a first step towards designing explanatory feedback by identifying that muscle tension could explain in part why some mental commands are incorrectly decoded and by showing that BCI users can deal with multimodal and multidimensional feedback.

In terms of perspective, we still have a lot to understand about BCI user training, and it is still necessary to redefine BCI user training approaches. A promising direction to change them in a relevant way, would be to make them satisfy principles and guidelines from human learning theories and educational science. At the level of the training tasks we propose our BCI users, educational science recommends to provide varied, adaptive and adapted training tasks. This raises a number of currently unanswered questions. How to design varied and relevant training tasks? What should these tasks train? In order to design adapted training tasks, we should also find out about how the user's profile impact BCI learning and performances. As mentioned earlier, we and others have conducted recent research going into that direction and which are worth being further studied [120, 188, 136, 139]. To design adaptive training tasks also requires to adapt the training tasks sequence to each user, overtime. How to do so to ensure an efficient and effective learning? The BCI community could learn on this topic from the field of Intelligent Tutoring Systems (ITS), which are tools specifically designed for digital education, to find an optimal sequence of training exercises for each user, depending of this user's skills, traits and states [189, 190]. It is thus also necessary to be able to quantify in a more refined way, beyond current classification-based metrics [191], what is a BCI control skill, in order to know what the user still has to learn. The PhD thesis of Jelena Mladenovic, who started just this year, aims specifically at exploring this direction of adaptive BCI training and operation.

At the level of the feedback provided, educational science recommends to exploit multimodal feedback and to provide a feedback that is explanatory rather than purely corrective, which it is so far for BCI. This raises the question of whether we can exploit other feedback modalities (audio, tactile) for training, and how? As mentioned before, recent results show that complementary tactile [182] or proprioceptive [192] feedback can enhance motor imagery BCI performances for instance, which confirms this is a promising direction to explore. Designing an explanatory feedback for BCI is currently very challenging given the lack of fundamental knowledge on motor imagery and on BCI feedback. For instance, why mental commands are sometimes erroneously recognized? Our work in [121] is only a very small step in that direction, which should be studied much deeper. A new PhD student coming at the end of this year will explore this area. Moreover, which feedback content providing to the user? Which feedback presentation should be used to represent this content? These are other crucial research questions that the BCI community will have to answer to design appropriate feedbacks for BCI, and thus to efficiently and effectively train our BCI users.

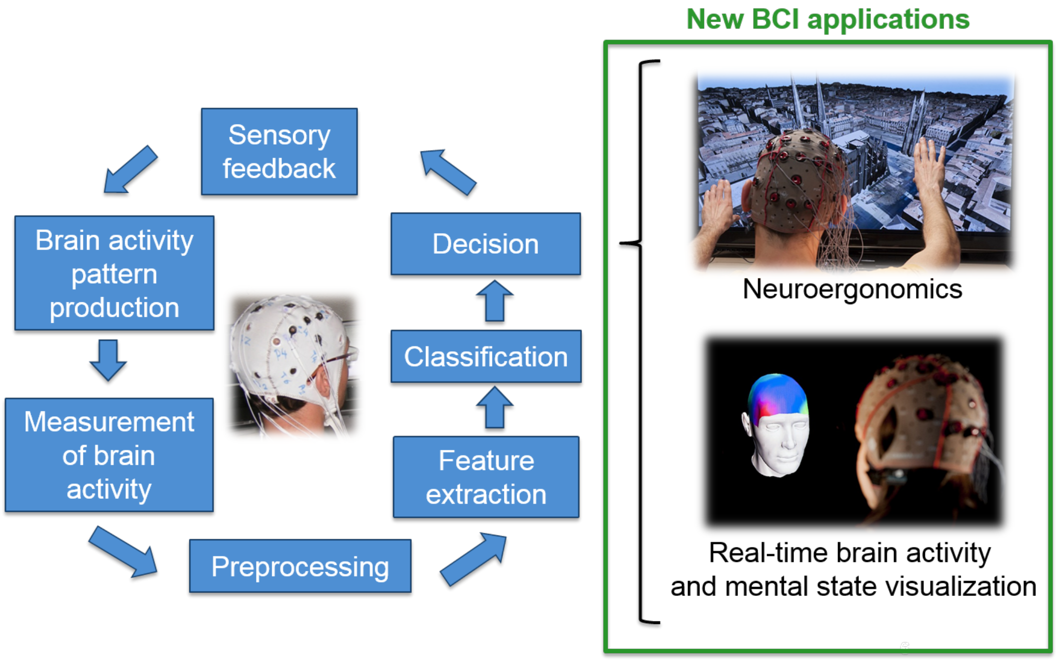
Finally, at the level of the training environment, improving BCI training requires to design motivating and engaging training environment. How to do so? How to keep BCI users being motivated and engaged in the training? Some recent works showed the positive impact of video games and virtual reality on BCI training and performance [193, 194]. There is now a need to understand why it is so, and to formalize these approaches to ensure the next generation of BCI training approaches will be motivating at all times, for all users.

From a non-technical and non-scientific point of view, all this research also depends on how to report such studies on BCI user training. Indeed, many different training tasks, feedbacks and environments could be tested and explored. Not all of them will lead to improve BCI training efficiency nor improved BCI performances. However, it is essential to know what works and what does not work to deepen our knowledge on BCI training. It would also be really inefficient, especially considering how costly (in time and money) BCI experiments on user training can be, if several research groups were to try the same experiments without knowing that other groups have already tried before but that it failed. This all points to the necessity to publish negative results in BCI research, to ensure an efficient research and to ensure access to all available relevant knowledge.

This chapter and the previous one focused on improving mental-imagery based BCI for communication and control applications. Indeed, their current usability, notably in terms of effectiveness (classification accuracy, robustness to noise) and efficiency (calibration time, user training time), prevents them from being widely used outside laboratories. However, current BCI technologies are potentially already mature and usable enough for other applications than communication and control. In the next chapter, we present our work towards using BCI technologies for such other applications, in particular for neuroergonomics and real-time brain activity and mental state visualization.

Chapter 3

Exploring other usages of BCI Technologies



Selected related Publications:

- J. Frey, M. Hachet, F. Lotte, *EEG-based Neuroergonomics for 3D User Interfaces: opportunities and challenges*, Le Travail Humain, accepted, 2016
- J Frey, M Daniel, J Castet, M Hachet, F Lotte, *Framework for Electroencephalography-based Evaluation of User Experience*, ACM SIGCHI Conference on Human Factors in Computing Systems (ACM CHI), pp. 2283-2294, 2016

- R. Gervais, J. Frey, A. Gay, F. Lotte, M. Hachet, *TOBE: Tangible Out-of-Body Experience*, Tangible, Embedded and Embodied Interaction (TEI), pp. 227-235, 2016
- J. Frey, A. Appriou, F. Lotte, M. Hachet, *Classifying EEG Signals during Stereoscopic Visualization to Estimate Visual Comfort*, Computational Intelligence and Neuroscience, Article ID 2758103, 2016
- D. Wobrock, J. Frey, D. Graeff, J.-B. de la Rivière, J. Castet, F. Lotte, *Continuous Mental Effort Evaluation during 3D Object Manipulation Tasks based on Brain and Physiological Signals*, Human-Computer Interaction - INTERACT, pp 472-487, 2015
- J. Frey, R. Gervais, S. Fleck, F. Lotte, M. Hachet, *Teegi: Tangible EEG Interface*, ACM User Interface Software and Technology (UIST) symposium, pp. 301-308, 2014
- C. Mühl, C. Jeunet, F. Lotte, *EEG-based Workload Estimation Across Affective Contexts*, Frontiers in Neurosciences section Neuroprosthetics, vol 8, no. 114, 2014
- J. Mercier-Ganady, F. Lotte, E. Loup-Escande, M. Marchal, A. Lécuyer, *The Mind-Mirror: See your Brain in Action in your head Using EEG and Augmented Reality*, IEEE Virtual Reality conference (VR 2014), pp. 33-38, 2014
- J. Frey, C. Mühl, F. Lotte, M. Hachet. *Review of the Use of Electroencephalography as an Evaluation Method for Human-Computer Interaction*, International Conference on Physiological Computing Systems (PhyCS 2014), pp. 214-223, 2014
- J. van Erp, F. Lotte, M. Tangermann, *Brain-Computer Interfaces: Beyond Medical Applications*, IEEE Computer, vol. 45, no. 4, pp. 26-34, 2012

Scientists that I (co-)supervised for this work:

- PhD student:
  - Jérémy Frey
- Post-doc:
  - Christian Mühl
- Master students:
  - Maxime Daniel
  - Dennis Wobrock
  - Léonard Pommereau
  - Camille Jeunet

Related research project:

- LIRA (LIfe-style Research Association) consortium - stress & relaxation project, 2011-2021
- DGA-DSTL Project "Assessing and Optimising Human-Machine Symbiosis through Neural signals for Big Data Analytics", 2014-2018 (co-Principal Investigator)

3.1 Introduction

While current BCIs are not yet usable enough to be widely used in practice for communication and control applications, they still have some unique properties that make them potentially already useful for other types of applications. Indeed, BCI technologies - i.e., the set of tools and methods used to design BCI - are the only technologies capable of estimating a person's mental state continuously, in real-time and in single trial. Even if this decoding is far from being perfect, this can still prove already useful for a number of applications for which speed and accuracy are less critical [14]. In particular we have explored two other usages of BCI technologies that can be already useful: neuroergonomics and real-time brain activity and mental state visualization.

Neuroergonomics refers to the use of knowledge and tools from neuroscience in order to assess the ergonomics qualities of various User Interfaces (UI) such as Human-Computer Interfaces (HCI) [195, 196, 197]. EEG-based BCI technologies can offer a unique contribution to that field by estimating continuously and in single trial some mental states of a user during interaction, and by using them as objective metrics to assess the User eXperience (UX) [198, 199]. We will show in this chapter that this is indeed feasible. Real-time brain activity and mental state visualization can enable anyone to see his/her own brain activity, understand better how this works and become aware of his/her own mental states. As we will show in this chapter, this can have multiple promising applications in domains as diverse as scientific mediation, education, neurofeedback, diagnostic and many others.

This chapter is organized as follows. Section 3.2 first present our contributions about using BCI technologies for neuroergonomics. More specifically, Section 3.2.2 presents our work on a neuroergonomic study of input devices and in particular on estimating the user's level of mental workload while he/she is interacting with an application, while section 3.2.3 is focused on output devices, here stereoscopic displays. Section 3.2.4 proposes some discussion and perspective about EEG-based neuroergonomics. Then, Section 3.3 presents our work on real-time brain activity and mental state visualization. Finally, the Chapter ends in Section 3.4 by some brief discussions. It should be noted that a substantial part of the works presented in this chapter is based on the PhD thesis of Jérémy Frey that I co-supervised [200].

3.2 Neuroergonomics

3.2.1 BCI technologies for neuroergonomics

HCI are increasingly used in a number of applications including industrial design, education, art or entertainment [201, 202, 203, 204]. As such, HCI and interaction techniques can be used by many different users with many varying skills and profiles. Therefore, designing them requires adequate evaluation tools to ensure a good UX for most targeted users [205, 204, 199]. To do so, a number of evaluation methods has been developed including behavioral studies, testbeds, questionnaires and inquiries, among others [206, 203, 204, 199]. This resulted in the design of more relevant, efficient and easy-to-use HCI.

Nevertheless, there is still a lot of room for improvements in the currently used evaluation methods. In particular, traditional evaluation methods could either be ambiguous, lack real-time recordings, or disrupt the interaction [206]. For instance, although behavioral studies are able to account in real-time for users' interactions, they can be hard to interpret since measures may not be specific, e.g., a high reaction time can be caused either by a low concentration level or a high mental workload [207]. Questionnaires and other inquiry-based methods such as "think aloud" and focus group all suffer from the same limitation: resulting measures are prone to be contaminated by ambiguities [208], social pressure [209] or participants' memory limitations [210]. There is therefore a need for more objective and continuous measures of the usability qualities of HCI that do not interrupt the user during interaction.

In order to obtain such measures of the user's inner-state during interaction, a recent promising research direction is to measure such states based on brain signals acquired from the user during interaction [206]. In other words, HCI could be evaluated by following a Neuroergonomics approach [196]. In particular EEG-based neuroergonomics seems very promising, since EEG are both accessible, portable, non-invasive and provide a high temporal resolution. Together with the available of efficient EEG signal processing tools developed for BCI technologies, this makes EEG suitable to measure a number of cognitive states that are relevant to assess HCI.

We notably conducted a survey of the literature (see [206] for details), which revealed that there are a number of mental states that are relevant to assess UX and that can be potentially estimated from EEG signals, at least to some extent and in laboratories conditions. In brief, the relevant mental states that could potentially be estimated from EEG include workload [211, 212, 213, 214], attention, vigilance and fatigue [215, 216, 211, 212, 217, 218, 219], error recognition [220, 221, 222], emotions [43], engagement, flow and immersion [211, 223, 224].

Altogether, measuring UX from EEG signals seems thus to be promising and feasible. In the following, we will see that analyzing EEG signals during HCI tasks using BCI technologies can provide us with information about the ergonomics of both the input devices and interaction techniques (Section 3.2.2), and the output, e.g., the display (Section 3.2.3).

3.2.2 Input devices: EEG-based workload estimation

A useful UX measure is the user's mental workload, i.e., the pressure on the user's working memory. Indeed, ideally, a UI should cause rather little mental workload to its users, to ensure the UI is cognitively easy to use, and to enable the user to devote his/her cognitive resources to the task rather than to the UI. Mental workload is typically measured using the NASA-TLX (Task-Load indeX) post-hoc questionnaire [207]. Even though it can be used to assess users' preferences regarding UI [225], NASA-TLX being a post-experiment measure, this is only a subjective and global measure that cannot inform on where and when the user experienced higher or lower workload. It therefore seems relevant to try to estimate mental workload from EEG signals, in order to obtain a continuous, objective and non-interrupting measure.

Interestingly enough, some works have started to use brain signals based measures of workload to compare 2D visual information displays [226, 227]. However, to the best of our knowledge, estimating mental workload from brain signals has never been explored to evaluate complex HCI such as 3DUI, although it could provide relevant evaluation metrics to complement the already used ones. 3D interaction tasks are more complex than 2D visualization for the user since 1) the user is actively interacting with the application, and not as passively observing it, and as such should decide what to do and how to do so, and 2) perceiving and interacting with a 3D environment is more cognitively demanding, since it required the user to perform 3D mental rotation tasks to successfully manipulate 3D objects or to orientate him/herself in the 3D environment. Therefore, as compared to existing works which only explored passive 2D visualizations, monitoring mental workload seems more relevant during 3D manipulation tasks, since the user is more likely to experience pressure on his/her cognitive resources. Moreover, most of the works on workload estimation from EEG signals, were lab based evaluations [213, 211, 228], using a very controlled approach. This naturally limits the range of real-world conditions they reflect. While this control is necessary to ensure the psychophysiological validity of the mental state detection, their results lack a certain ecological validity, and may not generalize to other contexts.

Therefore, we first tried to design an EEG-based workload estimator, using BCI tools, and studied how it performed across different contexts, in order to find out whether this estimator could be used later to assess the ergonomics qualities of real HCI or 3DUI. We then refined this estimator and explored it to discriminate mental workload levels in complex interaction tasks, notably during 3D object manipulation tasks, as well as during navigation tasks in a video game. In this last section, we also used it to compare the mental workload induced by different interaction devices.

EEG-based workload estimation across contexts

To be able to use EEG-based workload estimation in practice, it is necessary to ensure the workload estimator can work well across different contexts. Specifically, the influence of contextual changes of mental states (e.g., mood or motivation) that are predominant in the context of application have to be explored. That is why we tested the robustness of a workload estimator to the influence of (psychosocial) stress [58]. We let participants work under different levels of workload, while either under the impression of being observed and validated, or while being relaxed and free from this kind of pressure. We are interested in the effect of the contextual manipulation of stress on the workload estimator performance.

Protocol 12 female and 12 male participants were recruited for our experiment (mean age of 24.7 ± 7.9). We recorded EEG with 28 active electrodes in a 10/20 system without T7, T8, Fp1 and Fp2. In order to counterbalance the conditions (to avoid any order effect), we set up four scenarios composed of two blocks each. Therefore, we randomly begin with either relaxation or stress induction, and we randomly start with either low workload (0) or high workload (2).

In order to induce stress, we setup stress-induction protocol based on the Trier Social Stress Task (TSST) [229], which is a validated protocol. The stress induction is composed of three parts and requires the participation of three people, "the committee", who are presented as being body language experts. In the first part, a member of the committee asks the subject to prepare, during five minutes, a fake job interview for a teacher position. During the second part, the committee asks the person to do this job interview and to speak about himself for 5 minutes. They tell the subject that he is filmed (for a future behavioral analysis) and they take notes. The committee acts as being serious and neutral/unresponsive towards the subject. The third part is a 3 minute long arithmetic task (the subject has to count from 2083 to 0 by steps of 13) and to begin again at any mistake or hesitation. At the end of this protocol, in order to keep the stress level high, the committee tells the subject he will be filmed during the workload tasks and that he will have to do an another interview, which will be longer, and a self-evaluation after it. Furthermore, during the experiment, participants have to perform cognitive tasks, followed by feedbacks corresponding to their performance. During the stress condition, these feedbacks have been modified to display a performance 5 to 10 % lower than the real one. Thus, this protocol includes psychosocial stress and uncontrollability in order to maximize the chances to trigger a stress response for all the participants [230]. On the other hand, the goal of the relaxation induction was to create a condition in which participants would be able to relax. Thus, they were allowed to chose between resting in silence or select music/videos that make them feel calm [231].

In order to induce different level of mental workload, we used the N-back task [232]. With the N-back task, sixty white letters appear on screen the ones after the others (1500ms of break between them) on a black background. Among these letters, 25% are targets. When a letter appears the subject is asked to do a left click if this is a target letter, and a right click otherwise. Thus, for the 0-back task, which is our low workload condition, the target is the letter “X”: each time an “X” appears, the subject has to do a left click, and in all the other cases he has to do a right click. For the 2-back task, our high workload condition, the subject has to do a left click when the letter that appears is the same than two letters before (for example, if the sequence "C A C" appeared, the second "C" would be a target). Using behavioral measures, subjective questionnaires, and physiological signals we validated that this protocol indeed induced varying level of mental workload and varying levels of stress, see [119] for details.

Evaluation In terms of signal processing to built our workload estimator, we used the Filter Bank CSP (FBCSP) method [233]. The FBCSP method first consists in filtering each training EEG trial into multiple frequency bands using a bank of band-pass filters. Then for each of these bands, we optimized CSP filters. Then, the power of the spectrally and spatially filtered EEG signals over 2 sec is used as features, from which the most relevant ones are selected using the maximum Relevance Minimum Redundancy (mRMR) feature selection algorithm [234]. This amounts to selecting the most relevant pairs of spectral and spatial filters. Finally, the selected power features are used to train a shrinkage Linear Discriminant Analysis (LDA) classifier [112, 212] to discriminate low workload EEG trials from high workload ones.

The performance of our workload-level estimator was assessed using 6-fold stratified Cross-Validation (CV), separately for each subject. The goal of our work is to design a generic workload-level estimator, usable in practice, i.e., that can work across different affective contexts (here, different psychosocial stress levels). To do so, we performed different evaluations to estimate 1) the general performance of our system, independently of the affective context; 2) how it behaves *within* a given affective context and 3) how it behaves *across* different affective contexts, i.e, can a workload-level estimator calibrated on data from a given affective context (e.g., a relaxed condition) can be used to estimate workload in another affective context (e.g., a stressful condition)? Different sub-parts of the data were thus used for training and testing within our CV scheme for each condition, see [58] for details.

Results For the training and testing on the basis of all available data, those trials recorded during stress *and* non-stress context, we achieved an average classification accuracy of 76.1%, with performances between 58.7% and 95.4%, which is better than chance performances [138].

To evaluate the effects of testing in dependence of training context, we conducted a 2 (training context: no-stress, stress) × 2 (testing context: same-as-training, different-from-training) ANOVA. The main effect found for the testing context ($F(1,20) = 5.610$, $p = 0.028$, ${\eta_p}^2 = 0.219$) shows that the transfer from one context to another is problematic and results in a decrease of classifier performance (mean $= 69.4 \pm 9.7\%$) compared to testing on the same context as for the training (mean $= 72.4 \pm 9.4\%$).

Discussion and Conclusion We showed that workload can be classified on the basis of 2 seconds of EEG with an accuracy of 76.1%. This is comparable to previously reported results for such short intervals of data [214, 213]. Regarding the classifier generalization to different affective contexts, we show that a classifier created in a non-stressful context can generalize to a stressful context and vice versa. However, the training context has a significant influence on the classification performance, with decreasing performance for cross-context classification (i.e., from 72.4% to 69.4%). More generally spoken, the results suggest that the classification performance is not only dependent on quantitative factors, such as the numbers of channels, amount of training data, or length of trials, but also on qualitative factors, such as the affective context. This underlines the need for studies that identify such contextual factors and that elucidate ways to deal with detrimental effects related to their influence. From a practical point of view, this means that if we want to use such a workload estimator designed on EEG data collected during N-back tasks, to estimate workload during HCI tasks such as 3DUI, which is a widely different context, the estimator is likely to perform very poorly. By refining the design of this estimator, we will show in the following section that this can still be done.

Workload estimation during continuous 3D manipulation tasks

With the objective of moving from laboratory tasks (N-back tasks) to real HCI tasks for UX evaluation, we propose to assess users' mental workload during 3D object manipulation tasks (see Figure 3.1), based on EEG signals. We notably proposed a method to estimate workload levels from EEG in such a complex context, and studied workload levels during a 3D docking task [235].

Workload estimation To measure workload during 3D manipulations tasks, we first need a workload estimator. To obtain calibration EEG data, we used the N-back tasks described previously. We thus had our users alternate between "easy" tasks with the 0-back task and "difficult" blocks with the 2-back task. From the EEG data collected during these N-back tasks, we could build

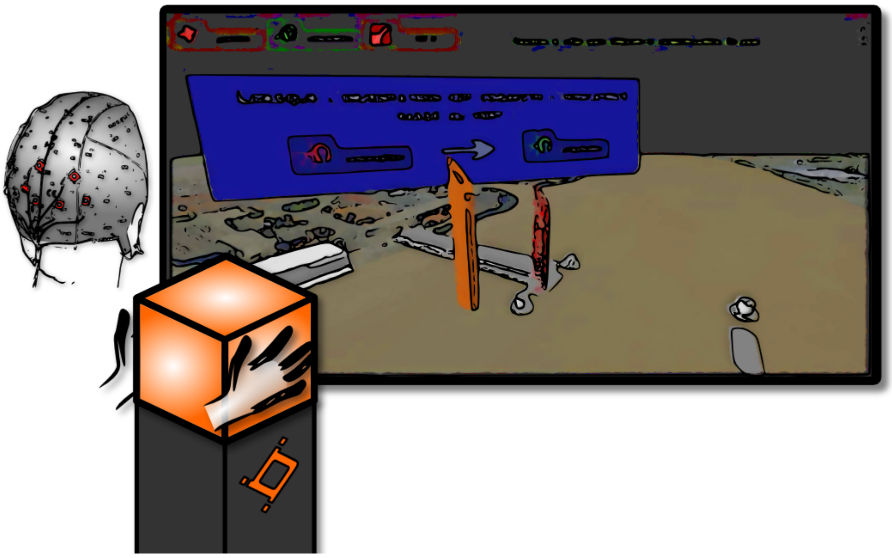


Figure 3.1: Schematic view of a user performing 3D manipulations tasks with the CubTile input device. His/her mental workload level is monitored based on EEG signals.

a workload classifier, using the FBCSP and an LDA classifier as described previously.

This signal processing approach is the one we used to discriminate workload levels from EEG signals between 0-back and 2-back tasks, i.e., within the same context on which the workload estimator was calibrated. However, as discussed before, EEG signals change between different contexts, which means that a workload estimator calibrated in a given context will have poorer performances when applied to a different context [58]. In our experiment, the final application context, i.e., 3D objects manipulation, is very different from the calibration context, i.e., the N-back tasks. Indeed, during the N-back tasks the user is moving very little as he/she is only performing mouse clicks, and exposed to very little visual stimulations as the N-back task only involves the display of white letters on a black background. On the contrary, manipulating 3D objects means that the user will be moving more and would be exposed to very rich visual stimulations. As such, a workload estimator simply calibrated on the N-back tasks and applied on the 3D object manipulation tasks is very likely to give very poor results or even to fail. Therefore, we modified the above mentioned signal processing approach to make it robust to EEG signal changes between the two contexts. In particular, rather than using basic CSP spatial filters, we used regularized CSP spatial filters [89] (described in Chapter 1, Section 1.4.1), that are robust to changes between calibration and use contexts. To do so, based on [236], we estimated the EEG signal covariance matrix from the calibration context (N-back tasks) and from the use context [1] (3D object manipulation tasks), and computed the Principal Components (PC) of the difference between these two matrices. These PC represent the directions along which EEG signals change between calibration and use. These directions are then used to regularize the CSP spatial filters as in [236], to ensure that the obtained spatial filters are invariant to these EEG signals changes.

We then used a shrinkage Linear Discriminant Analysis (sLDA) classifier [237] to learn which feature values correspond to a high or low workload level.

Evaluation during 3D manipulation tasks Once we have a classifier that can estimate workload levels from brain signals, we can use it to study mental effort during 3D objects manipulation tasks. With this objective in mind, we designed an experiment in which participants had to manipulate 3D objects using an interaction device known as the CubTile [238]. The CubTile is made of 5 orthogonal touch surfaces (Figure 3.1). Thanks to those many planes, the CubTile enables users to manipulate 3D objects with more degrees of freedom than with traditional input devices; they can easily choose the axis onto which they perform scaling (with a “pinch” movement), rotation (when at least 2 fingers draw such motion) or translation (when fingers slide on one of the surfaces). In our experiment, participants had to perform 3D docking tasks with the CubTile in order to build a bridge in 3D by assembling together its different parts (see Figure 3.1), while their EEG was recorded.

8 participants (2 females, age from 16 to 29) took part in this study. Each participant first took part to a series of N-back tasks to obtain calibration data to setup the workload classifier. The participant then had to construct the 3D bridge by assembling the bridge parts (e.g., the 4 supporting pillars and the road) one by one using the CubTile. In particular, the participant had to perform docking tasks, by translating, rotating and scaling the bridge parts, in order to put them at the correct location. The correct location was indicated to the user with proper 3D feedback, integrated to the 3D scene, in the form of text and color indicating how close he/she was from the correct position, scale and orientation. All the translations, rotations and scaling were controlled by the CubTile. The participant had to perform a set of 7 docking tasks, while their EEG were recorded. In particular, during the second task a time constraint was added and during the fifth task the controls were inverted on purpose, in order to test if we could measure a difference within physiological signals that would concur with such increases of difficulty (See [235] for details).

1. Note that this is only possible here because we perform an offline evaluation, after the 3D manipulation tasks have been performed and the corresponding EEG signals collected. It would not be possible to use the exact same algorithm for a real-time estimation of workload during 3D objects manipulation tasks as the covariance matrix of EEG signals during these tasks is not yet fully known.

Results For each user, we first setup a workload level classifier based on the signals collected during the calibration session (N-back tasks). Results revealed that the classifier could distinguish low workload from high workload during the N-back tasks with an average accuracy of 88.6%. By using this workload level classifiers obtained offline, we were able to estimate whether the current 2-seconds long time window of signals corresponds to a low or high workload for the user, during the manipulation tasks. This gave us unique insights into how much mental effort the participants were devoting to each task, and how these mental effort evolved over time.

We studied average workload levels, after normalization, to obtain a robust and reliable picture of the mental workload level associated to each task. A normalized workload index of +1 represents the highest mental workload participants had to endure while a normalized workload index of -1 denotes the lowest workload. Figure 3.2 displays the workload levels averaged over all participants and over the duration of each docking task.

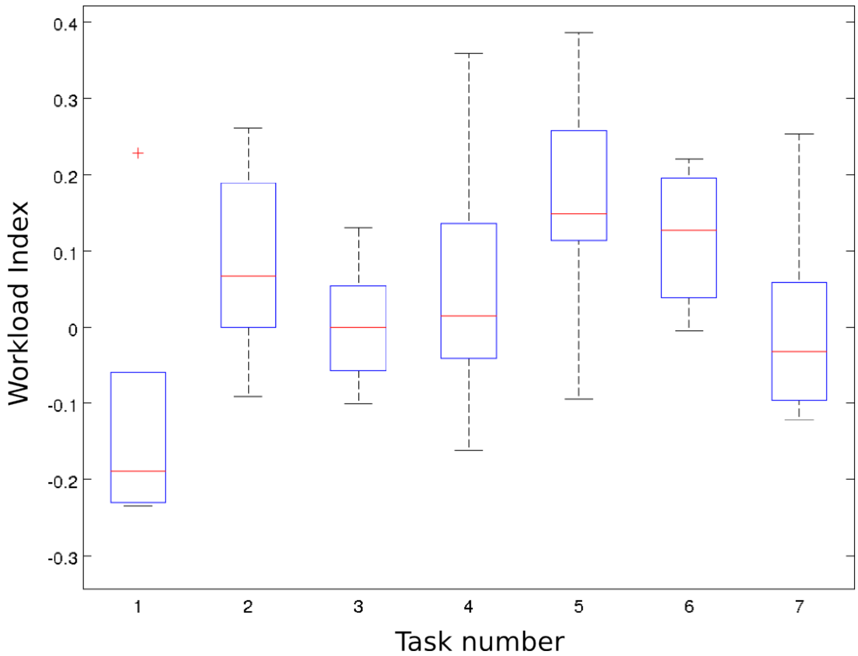


Figure 3.2: Average workload levels (averaged over participants and task duration) measured for the different 3D docking tasks.

A permutation test revealed that such workload levels distribution cannot happen by chance ($p < 0.001$ - see [235] for details), and thus that our workload classifier does find a workload level information during the 3D docking tasks that cannot be found by chance.

Discussion and Conclusion The observed workload levels suggest that despite the novelty and the complexity of the interaction – handling at the same time rotation, translation and scaling of elements in a 3D environment right from the beginning – the participants did not make an important mental effort to complete the first task. That could be due to the practicality of the CubTile, which may ease 3D interaction thanks to its additional degrees of freedom compared to a traditional input device such as a mouse.

When a constraint appeared concurrently with the second task – pillars were "falling" continuously from the sky and had to be positioned quickly before they touched the ground – the workload index increased substantially. This is consistent with the sudden pressure that was exerted on users. As one could expect, the mental workload lowered and settled in tasks 3 and 4, during which there was no more time pressure – but still more complex manipulations compared to task 1.

We purposely inverted the commands during the fifth task to disorientate participants. As a matter of fact, this is the moment when the workload index was the highest on average among all participants. Then, after this sudden surge of mental stress, once again the measured workload has been reduced in the two subsequent tasks. Interestingly enough, for task 6, in which the control commands were inverted back to normal, the workload indeed decreased as compared to that of task 5, but was still higher than for the other tasks. This probably reflects the fact that users had somehow integrated the counterintuitive manipulation technique and had to change again the gestures they used to manipulate the 3D object, thus being forced to forget what they had just learned in task 5 which resulted in a high workload.

Overall, the mental workload that was measured with EEG along the course of the interaction matches the design of the tasks. Workload increased when a sensitive element of the interaction was deprived – e.g. time or commands – which can be explained by the need to overcome what participants have learned previously and re-learn how to handle the new environment. Afterwards, when going back to the previous scheme, the workload goes back to a low level, as could be expected. Altogether, these results suggested that continuous mental workload monitoring was possible and could provide us with interesting insights about how cognitively easy-to-use a given 3D interaction technique can be. Such approach could thus become a relevant tool to complement existing 3DUI evaluation tools. To further validate its usefulness and relevance, it is however necessary to use it to assess other 3D interaction tasks such as navigation or application control, as well as to use it to compare different interaction devices or techniques. This is what we explored in the study presented in the next section.

Workload estimation during complex navigation tasks and interface comparisons

Based on the studies and tools presented above, we went a step further in order to validate them on another realistic task (see Figure 3.3) and to show that they can be used to compare different HCIs [239]. In particular we built a gamified 3D Virtual Environment (VE) to assess workload during ecological and realistic interaction tasks. Below we describe this VE and the neuroergonomics experiments conducted with it.

Figure 3.3: We demonstrate how EEG can be used to evaluate UX. For example, a keyboard (*left*) can be compared with a touch interface (*middle*) using a continuous measure of mental workload (*right*)

Virtual 3D maze The VE takes the form a maze where players have to learn and reproduce a path by triggering directions at regular intervals (see Figure 3.4). A character displayed with a third person perspective moves by itself at a predefined speed inside orthogonal tunnels. Soon after the character enters a new tunnel, symbols appear on-screen. Those symbols are basic 2D shapes, such as squares, circles, triangles, diamonds or stars, and their positions (bottom, top, left or right) indicate which directions are "opened". Players must select one of these symbols before the character reaches the end of an intersection, either by pressing a key or touching the screen. If users respond too early, i.e., before symbols appeared, too late, or if they select a direction that does not exist, they loose points and the character "dies" by smashing against a wall, respawning soon after at the beginning of the current tunnel.

The main element of the gameplay consists in selecting the directions in the correct order. Indeed, one level comprised two phases. During the “learning” phase a particular sequence of symbols is highlighted; at each symbols’ appearance one of them is bouncing to indicate the correct direction. Another cue takes the form of a “breadcrumb trail”, a beam of light that precedes the character and points to the correct direction (see Figure 3.4, middle). Selecting an available but incorrect direction does not result in the character’s “death” but leads to a loss of points. A visual feedback is given to users when they select a direction: the corresponding symbol turns green if the choice is correct and red otherwise. When the end of the maze is reached, the character

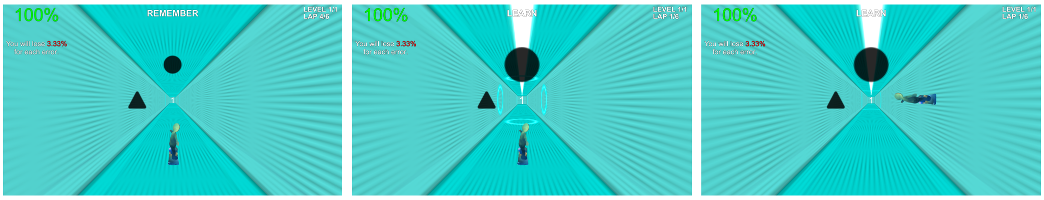


Figure 3.4: The virtual environment, where players control a character that moves by itself inside a 3D maze. *Left*: Symbols appear in each tunnel to indicate the possible directions for the next turn; players have to select a particular sequence of symbols/directions. *Middle*: During the "learning" phase, the correct direction is highlighted by a breadcrumb trail and the associated symbol bounces (here the disc on top). *Right*: Controls depend on the position of the character. If the character is on the right side, players have to press *right* in order to go *up*.

loops over the entire path so that players have another opportunity to learn the sequence. When the training phase ends, the "recall" phase follows. The symbols are identical but the cues are no more displayed; players have to remember by themselves the right path. Symbols position in each tunnel and symbols sequence are randomly drawn when a new level starts.

Beside learning a sequence, the principal challenge comes from *how* the directions are selected. The third-person view fulfills a purpose: the input device that users are controlling – i.e. keyboard or touch screen – is mapped to the *character position*. Since the character is a futuristic surfer that defies the law of gravity, it slides by itself from the bottom of the tunnel to one of the walls or to the ceiling from time to time. In this latter situation, when the character is upside down, commands are inverted compared to what players are used to, even though symbols remain in the same positions (See Figure 3.4, right).

Our VE possesses several characteristics that could be used to induce different levels of mental workload. We can notably adjust 4 parameters:

- *Maze depth*: the number of tunnels players have to cross before reaching the end of the maze, hence the length of the symbols sequence they have to learn. More symbols to be held in the working memory increases workload [214, 240].
- *Number of directions*: at each intersection, up to 4 directions are "opened" in the maze; the complexity of the symbols sequence grows as this number increases.
- *Game speed*: the pace of the game can be adjusted to increase temporal pressure. When the speed increases symbols appear sooner and users must respond quicker, thus increasing overall stress [207]. -
- *Spatial orientation*: in order to keep selecting the correct directions, users have to perform a mental rotation if the character they control jumps

from the floor to the walls or to the ceiling. Depending on the spatial abilities of users, this mechanism can cause an important cognitive load [241].

We used those mechanisms and dimensions to create 4 different difficulty levels for the game: "EASY", "MEDIUM", "HARD" and "ULTRA". These levels affect mostly (symbolic) memory load and time pressure. These difficulty levels were validated using the NASA-TLX questionnaire (see [239] for details).

To estimate workload levels during these 3D navigation tasks, we used the same calibration and signal processing approach described previously. In other words, we calibrated our workload classifier using EEG data collected during N-back tasks, with regularized FBCSP methods (regularized to ensure robustness across contexts) and a shrinkage LDA as classifier [239].

Evaluation The main study consisted in the evaluation of the game environment with two different types of interfaces using EEG recordings. In particular we compared navigating using a keyboard (KEYBOARD) or a touch screen (TOUCH). KEYBOARD uses *indirect* interactions (Figure 3.3, left). In due time, left, right, up or down arrow keys are used to send the character in the tunnel that is situated to *its* left, right, top or bottom. Indeed, we have seen previously that in our VE players have to orientate themselves depending of the position of the character. If the character is moving on the sides, players have to perform a mental rotation of 90° if it is on the ceiling then the angle is 180° i.e. commands are inverted. TOUCH uses *direct* interaction (Figure 3.3, middle). Usually, with touch screen, pointing is co-located with software events, since users can directly indicate where they want to interact. However, in our case, we decided to mimic exactly the behavior of the keyboard interface. That is to say that with the touch screen as well players have to orientate themselves depending on the position of the character. Hence, if the character is positioned on the *left*, players have to touch the *right* fringe of the screen in order to go *up*. This is mostly counter-intuitive since players have to inhibit the urge to point to the actual direction they want to go; there is a cognitive dissonance. Thus, we hypothesize that it will lead to an overall higher workload compared to the indirect interface (keyboard).

12 participants took part in this study (3 females), mean age 26.25 (SD: 3.70). Participants played the game using each one of the 2 interaction techniques (KEYBOARD or TOUCH) in a random order. The four levels of difficulty (EASY, MEDIUM, HARD, ULTRA) appeared twice with each interaction technique, in a random order.

Results Results of estimated workload levels across interaction techniques and workload levels are shown on Figure 3.5. The statistical analysis of the classifier output during the game session showed a significant effect of the difficulty factor ($p < 0.01$); the workload index increasing along the difficulty of the levels. The post-hoc analysis showed that all difficulty levels significantly differs one from the other with $p < 0.01$; except for the MEDIUM level, which differs from EASY with $p < 0.05$ and with HARD only by a margin ($p = 0.11$). There was a significant effect of the interaction factor as well ($p < 0.01$), the workload being higher on average during the TOUCH condition. There was no interaction between difficulty and interaction factors.

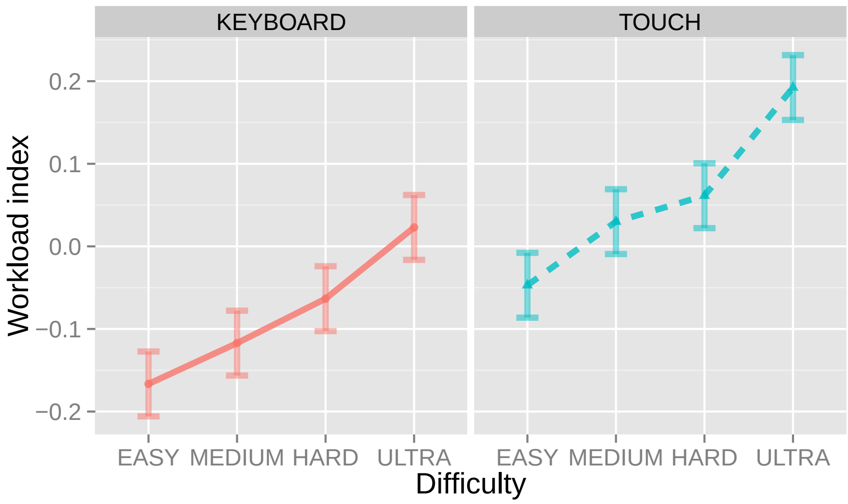


Figure 3.5: The workload index significantly differs across difficulties and between interaction techniques.

Discussion and Conclusion The workload index as computed with EEG showed significant differences that match the intended design of the difficulty levels. The workload index was also significantly higher in the TOUCH condition compared to the KEYBOARD condition. This thus validate our workload estimator as a useful tool to estimate workload due to the task content (here difficulty level) or the interaction technique (here KEYBOARD vs TOUCH). Interestingly enough, this workload estimator can provide insights that behavioral measures cannot. For instance, participants had similar performances at the task (finding the correct path in the Maze) with both KEYBOARD and TOUCH [239]. However, EEG analyzes revealed that the workload was significantly higher in the TOUCH condition, meaning that users had to allocate significantly more cognitive resources to reach the same performance.

Above all, an evaluation method based on EEG enables a continuous monitoring of users. The intended use case of our method is to enroll dedicated testers that would wear the EEG equipment and perform well during the calibration tasks. As a matter of fact, the best performer during workload calibration shows patterns that clearly meet the expectations concerning both difficulty levels and interactions, as pictured in Figure 3.6.

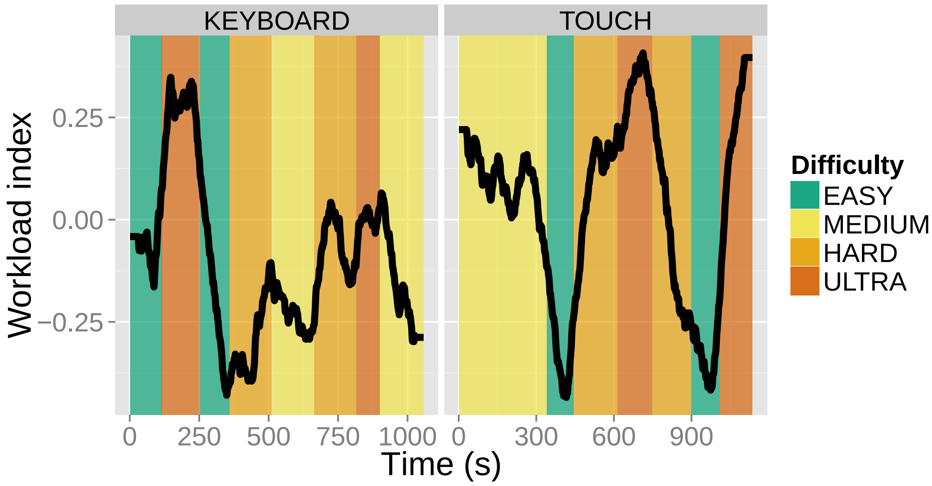


Figure 3.6: Workload index over time for our best participant – 60s smoothing window. *Left*: KEYBOARD condition, *right*: TOUCH condition. Background color represents the corresponding difficulty level.

3.2.3 Output devices: Estimating visual comfort during stereoscopic visualization in EEG signals

Interestingly enough, EEG-based BCI tools can also provide some insights on how the user perceives the application and its output devices, notably its displays. In particular, we showed how EEG-based neuroergonomics can be used to study stereoscopic displays [2] [242]. Stereoscopic displays have been developed and used for years, for example to improve data visualization or to better manipulate virtual objects [243], or more recently with "3D" movies or game consoles. Yet, irrespectively of the methods used to display stereoscopic images, visual discomfort or even pain could occur when the stereoscopic effect is too strong [244].

All stereoscopic displays use the same mechanism to give the illusion of depth. They send a different image to the left and right eyes. As with natural vision, the visual fields of our eyes overlap and the difference between the two images helps our brain to estimate objects' distance. To facilitate images merge, observers rely on two mechanisms. First, they need to maintain the point of interest at the same place on both their retinas. This is why the closer an object gets, the more eyeballs rotate inward. This is called "vergence", and it also happens with stereoscopic displays. Second, in a way similar to how camera lenses operate, crystalline lenses need to focus light beams. They deform accordingly to objects' position in order to obtain a clear picture. This other physiological phenomenon is called "accommodation" and

2. stereoscopic displays enable their users to perceive images and videos in 3D, to perceive the depth and volume of the displayed objects

is *not* replicated with stereoscopic displays. In a natural environment, vergence and accommodation are locked to objects' positions and occur altogether. But since the focal plane in stereoscopic displays is fixed, accommodation will not change. The discrepancy between vergence and accommodation is called the "vergence-accommodation conflict" (VAC). VAC is one of the major causes of the symptoms associated with visual discomfort and visual fatigue in stereoscopic displays [244]. While guidelines exists to limit VAC [245], they do not take into accounts screen settings, viewing angle or individual differences, making them difficult to use as is. In order to mitigate those symptoms and adapt the viewing experience to each user, we propose a method to discriminate uncomfortable situations from comfortable ones in EEG signals.

Protocol

We studied the appearance of virtual objects which were presented to participants at different apparent depths for a few seconds (see Figure 3.7). We created two conditions: objects appeared either at a comfortable position ("C" condition) or at an uncomfortable position ("NC" condition). We defined ranges inside and outside the zone of comfort according to [245]. Subjective questionnaires validated that these two conditions indeed significantly differed in the visual comfort they induced [242].

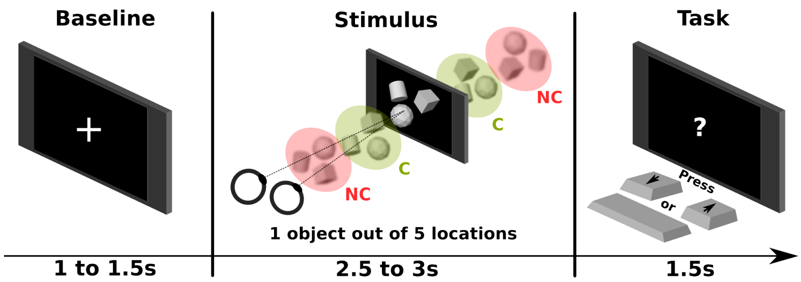


Figure 3.7: One trial: cross (baseline), object at random depth, task.

In order to assess participants capacity to situate virtual objects in space and to maintain their vigilance high during the whole experiment, they had to perform a task. When a question mark was shown on screen, "down", "space" or "up" keys were pressed to indicate whether objects appeared "in front of", "as far as" or "behind" the screen. A trial started with a neutral stimulus, a 2D cross appearing on-screen for a duration comprised between 1 and 1.5s (Figure 3.7). Then the virtual object appeared for 2.5 to 3s. Finally, a question mark appeared for 1.5s, a period during which participants had to perform the task. After that a new trial began. 12 participants took part in the experiment: 5 females, 7 males, mean age 22.33 (SD=1.15).

Analyzes

EEG were band-pass filtered in 0.5-25Hz. We extracted epochs of EEG around C and NC stimuli onsets, from -1s to +2.5s. Epochs containing muscle artefacts ($\approx 10\%$) were automatically rejected using the pop_autorej function from EEGLAB [246]. EOG activity was suppressed using the ADJUST toolbox 1.1. Averaged ERPs across participants indicated that ERPs had a higher positive peak in C (see Figure 3.8).

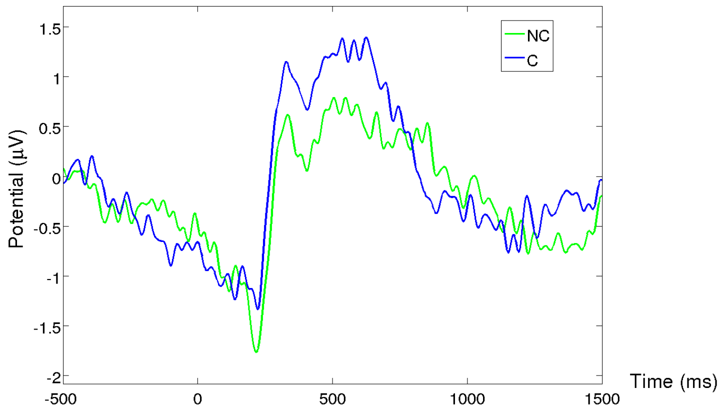


Figure 3.8: Average ERP across 28 EEG electrodes and 12 participants. Blue: comfort condition; green: no-comfort condition (≈ 160 trials each). The stereoscopic object appears at t=0ms.

For classifying those ERPs in single trial, we first split the EEG dataset of each participant in two. The first half of the trials was used as a training set, the second half was used as a testing set. Feature extraction relied on the spatial filter proposed by [247]. We used 5 such spatial filters and used a time window of 1s. In order to reduce the number of features, we decimated the signal by a factor 16. We used shrinkage LDA as a classifier, as recommended in [248] for ERP classification.

Although we used 1s time windows as a basis for our analyzes, we tested longer stimuli by clustering trials with Monte Carlo simulations. The principle is as follows: studying 3 presentations, we cluster 3 similar trials drawn from the testing set (e.g., “no-comfort”, 3xNC). Then we look at individual classification results from the system (e.g., NC-NC-C) and keep the label which has the majority – in this case NC, the resulting classification is correct for this cluster. Had the classifier labelled trials as “C-C-C”, the cluster would have been erroneously labeled as “C”. Different combinations of trials were drawn from the testing set to compute the scores for n=3,5,7.

Results

We were able to predict with an average classification accuracy of 63.30 $\pm$ 7.64% the visual comfort experienced by viewers (range [53.2 - 76.1]). With Monte Carlo simulations, we investigated how the system would perform with the appearance of several images from the same condition. Classifier accuracy reached 68.91 $\pm$ 10.32% over 3 trials. Over 5 trials the classification reached 90% for some users, resulting in a 71.83 $\pm$ 12.28% average. With n=7, one-third of the participants reached 90% or more (74.08 $\pm$ 13.39% on average).

Discussion and Conclusion

We described an innovative system that can distinguish uncomfortable stereoscopic viewing conditions from comfortable ones by relying on EEG signals. Using short time windows (features were extracted over 1s), we set the basis of a tool capable of monitoring user experience with stereoscopic displays in near real-time. Such a passive BCI can adapt the parameters to users' state (e.g., mental fatigue is likely to relate to visual fatigue) throughout the viewing. Moreover, a passive BCI does not disrupt work or the narrative of the stereoscopic environment. A passive stereoscopic comfort detector could potentially be useful for multiple applications, as a tool to: 1) objectively compare (possibly offline) different stereoscopic displays, 2) dynamically enhance stereoscopic effects, by increasing discrepancy without causing discomfort, 3) quickly calibrate stereoscopic displays, 4) dynamically adapt discrepancy to avoid discomfort (e.g., during 3D movies) or voluntarily cause discomfort (e.g., for basic science studies about perception), among many others.

3.2.4 Conclusion and perspectives on Neuroergonomics

So far in this chapter, we have shown that EEG-based BCI technologies can prove already useful for neuroergonomics. In particular we have shown how they can be used to estimate reliably mental workload levels from EEG signals, including during complex and demanding interaction tasks such a 3D manipulation or navigation tasks. This can be used to objectively quantify how cognitively demanding a given interaction task is and/or to compare different interaction devices or techniques. We have also shown how EEG-based BCI tools can be used to estimate visual comfort during stereoscopic visualization, which could provide additional insights on the HCI output devices.

These different works clearly showed the promising potential of a neuroergonomics approach based on EEG to evaluate HCI. However, research works in this direction are relatively recent, and there are still a number of scientific challenges that remain to be solved to be able to widen the applicability of this neuroergonomics approach in practice to the design of HCI. These challenges includes dealing with possible artifacts that may occurs with freely moving users interacting with a device, shortening or removing the calibration sessions required to induce the mental states of interest (e.g., the N-back tasks for mental workload) and build a classifier or studying from a neuroscientific point of view other cognitive states that could be relevant to study HCI, and how they are reflected in EEG signals, among many others. The interested reader can look into our recent paper in [249] for more open challenges and opportunities in EEG-based Neuroergonomics.

Moreover, all the work presented and discussed so far was focused on the use of EEG signals as the measure of brain activity. However as far as neuroergonomics is concerned other signals could and should be used to measure brain activity and/or assess the users' mental states. This notably include using other brain activity measure such as functional Near Infrared Spectroscopy (fNIRS), which is very promising for this domain, as well as physiological signals such as heart rate or galvanic skin response that could complement brain signals [250, 251, 252].

3.3 Real-time brain activity and mental state visualization

Another area in which BCI technologies can prove useful is the area of real-time brain activity and mental state visualization. Indeed, by combining BCI technologies with HCI approaches, it is possible to design engaging and easy-to-use tools that enable anyone to visualize in real-time his/her own brain activity and mental states or those of others, with many potential applications. We created three such systems, named respectively The Mind-Mirror, Teegi and Tobe. They are described hereafter.

3.3.1 The Mind-Mirror

The Mind-Mirror enables its users to visualize their own brain activity in real-time, in their own head [253]. It uses augmented reality [254] and head-tracking to overlay a representation of an active brain on top of the user's head, seen in a semi-reflective screen (see Figure 3.9). This gives the illusion that the user can see his/her own brain in his head, in a mirror, in activity, the EEG power from different channels being represented in real-time on the surface of the brain. The color of the whole brain can also be changed to inform about the mental state of the user, e.g., about the classifier output in a classical BCI experiment. The Mind-Mirror was tested as a feedback to train users to control an attention-based BCI in which the user had to perform concentration or relaxation tasks to control two different commands. The Mind-Mirror was compared to a classical gauge feedback for the same task. Results showed that users found the Mind-Mirror to be indeed an engaging and innovative visualization tool. In terms of performance however, it was not better than the classical gauge feedback. This suggests that selecting and visually enhancing the relevant information is probably necessary to ease the user perception and understanding of his/her own brain activity, at least for neurofeedback scenario

Figure 3.9: the Mind Mirror: See your brain in action in your own head.

3.3.2 Teegi

Teegi is a tangible EEG interface, designed to enable users that are naive with EEG to get to know more about how EEG works, in an accessible and engaging way [255]. As with the Mind-Mirror, Teegi enables users to visualize their own brain activity in real-time (See Figure 3.10, Left). Contrary to the former, Teegi is based on educational science principles to ease the user understanding of brain activity [256, 257]. In particular, it is based on a tangible system, the user's brain activity being projected on the head of a puppet (with a friendly and humanoid appearance to enhance attractiveness), that can be easily manipulated, to favor exploration [256, 257, 258]. This puppet is tracked, which allows us to co-locate the projection with the puppet's head, at any time. Hence, the user can easily visualize a realistic modeling of the EEG signals in any part of the scalp by manipulating the puppet, while maintaining a good spatial topology of the observed data.

EEG signals can also be filtered to reveal some specific EEG phenomenons (e.g., sensorimotor rhythms, occipital alpha rhythm), hence enhancing the relevant EEG signals. More precisely, three different filters can be applied to the raw EEG data enabling users to investigate influences of motor motions,

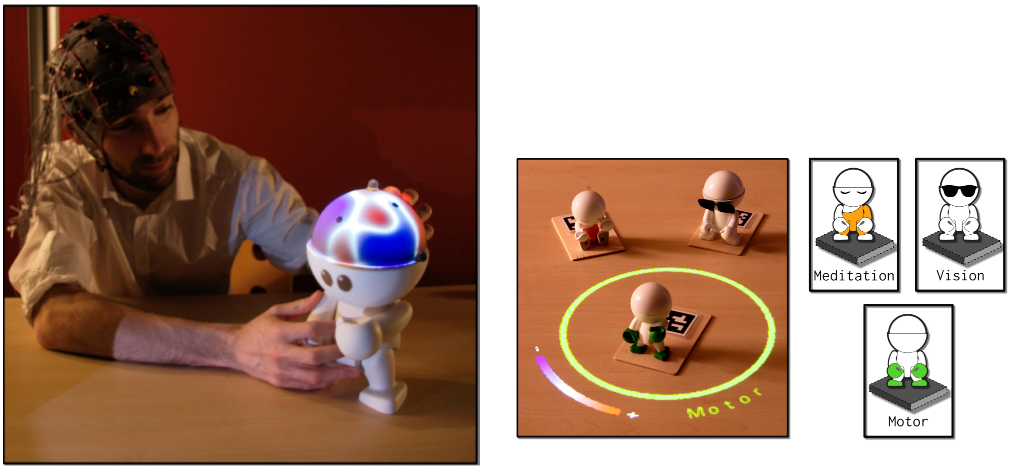


Figure 3.10: Left: Teegi, a Tangible EEG Interface. Right: the Mini-Teegi that can be placed in a circle on the table to apply some filters to reveal specific brain phenomenons

visual activities or meditation, on their brain activity in real-time. To remain consistent with the tangible philosophy of this project, we decided to control the filters by way of small tangible characters (mini-Teegis - see Figure 3.10, right) that can be moved on a "filter area", which is highlighted on the table by a projected halo. For example, if a user wants to apply a filter that will allow her to better see what happens when moving her hand, she just needs to take the dedicated mini-Teegi, i.e. the one with the colored hands, and to move it to the filter area. Then, by moving her right hand, she should see changes in EEG amplitude on the left hemisphere of Teegi's head, notably post-movement Beta ERS in the C3 electrode area.

Teegi was so far explored only for scientific outreach and education, to help users to know more about the brain and EEG. A small user study with 10 participants, revealed that those participants knew more about EEG and the brain (e.g., about the functions of some brain areas) after having used Teegi. Questionnaires filled in by participants also shown that the interaction with Teegi was very enjoyable and motivating. See [255] for details.

3.3.3 Tobe

Tobe (to be pronounced ['tobi']), is a Tangible Out-of-Body Experience shaped as a tangible avatar (see Figure 3.11, Left) [259]. This avatar lets users freely explore and represent their physiological (ElectroCardioGraphy - ECG, Galvanic Skin Response - GSR, respiration, etc.) and brain (EEG) signals, displayed on the avatar itself using Spatial Augmented Reality (SAR) [260]. SAR adds dynamic graphics to real-world surfaces using projected light. The overarching goal of Tobe is to help one reflect on his/her physiological and mental states in *his/her* own way. The main activity would be for users to actively build from the ground up their own self-representation and then visualize physiological signals and mental states through it. As such, we designed a modular toolkit around Tobe that can be used to customize any part of the system. As compared to Teegi, Tobe thus enables to also visualize physiological signals that are not of brain origin, to visualize high-level mental states (workload, arousal, valence, etc.) and to customize the whole system, so that each user can visualize the signals and states he/she wants to, in the way he/she wants to. In this sense, Teegi can be seen as one instance, one possible implementation made with the Tobe framework. To make Tobe attractive, clear and meaningful, its form factor and visualization were created with substantial users' input, through a co-design approach [259].

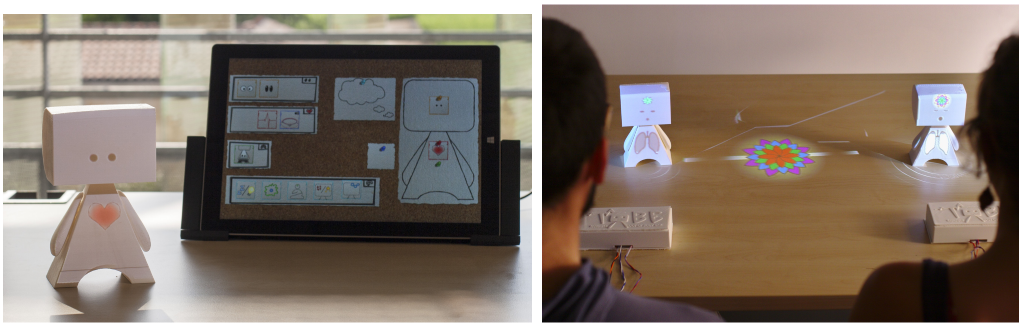


Figure 3.11: Left: Tobe, the tangible avatar displaying real-time physiological readings and high-level mental states along with the interface to control the different visualizations. Right: Multi-users application: relaxation through cardiac coherence.

As Teegi, Tobe is a tangible and anthropomorphic avatar, as we have found evidence in the literature that this combination of anthropomorphism and tangibility can foster social presence and likability [261, 258].

We investigated how everyday people picture their physiology and we validated the acceptability of Tobe in a scientific museum. We also tested Tobe as a relaxation device for two users (see Figure 3.11, Right). This version of Tobe relies only on respiration and heart rate variability. It relates to cardiac coherence: when someone takes deep breaths, slowly ($\approx$ 10s periods) and regularly, her or his heart rate (HR) varies accordingly and the resulting state has positive impact on well-being [262]. 14 participants joined our experiment, by pairs, during which they were instructed to synchronize their *hearts*. In order to do so, they had to both reach cardiac coherence while breathing on the same rhythm – with no other way to communicate than using their Tobes. Indeed, they could only see each other Tobes'. Qualitative results suggest that Tobe could be employed as a proxy for interpersonal communications and that it has an interesting potential for enhancing well-being (see [259] for details). Although this last application is not BCI nor EEG, it shows that showing users they own physiological signals or states can lead to promising applications. Indeed, as we will see in the next section, there are a number of applications for which real-time visualization of brain signals or mental states can be promising.

3.3.4 Potential applications

The different tools presented above are different ways to make users aware of their own brain activity, physiological signals and mental states in general. They were also designed to be accessible and attractive to anyone. As such, we can envision many potential applications of these technologies.

When considering a single user, those tools can be used as a biofeedback device with a specific goal, e.g. to reduce stress or increase attention (as we did with the MindMirror [253]), or to gain knowledge about one self. A feedback about workload and vigilance would prevent overwork. Insights gathered from an introspection session with them could also be employed to *act* better. For example, it might be useful to realize that you are irritated before answering harshly to beloved ones.

Education and scientific mediation is another promising area, to explain lay person how the brain, EEG and BCIs work, as we have started to do with Teegi [255]. Indeed, some studies have identified that misconceptions about brain functions prevail in the general public, hence stressing the need for tools to educate people about it [263, 264]. Moreover, people/learners should investigate and manipulate in order to become conscious of complex phenomena, change their misconceptions and construct scientific knowledge [256]. Our tangible tools revealing brain activity and mental states enable people to do exactly that.

Still considering a single user, but potentially one or multiple observers, our tools could be used in a medical context. Indeed, in stroke rehabilitation, patients with motor disabilities may regain mobility after long and difficult sessions of reeducation. However, occasional drawbacks may create anxiety and a counterproductive attitude towards therapy, which leads to even more anxiety. Our tools could help patients and therapists acknowledge this affective state and break this vicious circle. Autistic persons could also benefit from using Tobe since it is difficult for them and their relatives to gauge their inner state. Explicit arousal could help their integration into society. An offline experiment – i.e. after signals were recorded – pointed to this direction [265].

Considering multiple users, our tools could be used as an alternate communication channel during casual interactions and would help to explore connections with relatives, discover and learn from strangers or improve collaboration and efficiency with coworkers. They could also summarize the state of a group. A real-time feedback from the audience would be a valuable tool for every speaker or performer. To pace a course, a teacher could use one Tobe as an overall index that aggregates the attention level of every student in the classroom. Through behavioral measures and with a feedback given afterward, this was investigated in [266].

Altogether such tools could be defined as *introspectibles*, giving anyone new way to perform introspection and interoception, with many potential applications [267].

3.4 Discussion and perspectives on BCI technologies usages

In this chapter we have explored alternative usages of BCI technologies, beyond communication and control only (which were the focus of the previous chapters). In particular we have explored using EEG-based BCI technologies for neuroergonomics, to assess the ergonomic qualities of various HCI such as 3DUI. We have notably shown that can we could estimate workload levels continuously during complex interaction tasks, as well as estimate visual comfort with stereoscopic displays. We have also explored using EEG-based BCI technologies for real-time visualization of brain and physiological signals and mental states. We have proposed three systems to do so, exploiting augmented reality and tangible interaction: the MindMirror, Teegi and Tobe. We have shown they could be potentially used for multiple applications such as education, scientific mediation, neurofeedback, stroke rehabilitation, interaction with autistic people, group interaction, interoception and many others.

These new usages of EEG-based BCI technologies thus open many interesting research directions and potential future works. Regarding EEG-based neuroergonomics, it would be necessary to explore other mental states that would be relevant to assess HCI and that could be estimated in EEG signals. This would enrich the repertoire of mental states and insights that could be provided by EEG, thus complementing efficiently existing methods. Overall, we should work towards building a complete EEG-based evaluation framework, with dedicated calibration and mental state induction protocols, signal processing tools, and evaluation protocols to assess such states during actual HCI tasks. Our work in [239] goes into that direction, but much remains to be done. We are notably continuing our work in this area with the Immersion company, a Bordeaux-based company designing and selling 3DUI technologies (http://www.immersion.fr/).

Regarding real-time brain activity and mental states visualization - the so-called *introspectibles*, we have so far established proofs of concept and potential applications. But we still have to actually design, validate and use such systems in practice for these actual applications. Currently, we are targeting using them for well-being applications and in schools for education.

Discussion and Perspectives

Summary of the contributions

The works presented in this manuscript all aimed at addressing the usability issue of EEG-based BCIs, which currently prevents them from being widely used in practice outside laboratories. Indeed, as mentioned multiple times, EEG-based BCIs are not yet efficient - they are slow and require long calibration and training times - nor effective - they often recognize an erroneous mental command from the user. Our work from these last years have targeted this lack of usability at three different levels: 1) the EEG signal processing level, 2) the human learning level and 3) the application and usage level.

At the signal processing level, we have proposed new methods to increase BCIs effectiveness, in particular to improve the decoding accuracy of mental-imagery (MI)-based BCIs despite the noisy and non-stationary nature of EEG signals. To do so, we have explored alternative features, namely predictive complexity, multifractal cumulants, waveform length and phase locking values, and proposed optimal spatial filters for the last two of them. We have showed that when combined with classical band power features, such alternative features could all improve the overall BCI decoding accuracy. We have also proposed algorithms to robustly optimize CSP spatial filters, thanks to various regularization functions or robust averaging in Riemannian space, which also led to improved classification accuracy. Finally, to improve EEG-based BCIs efficiency at the signal processing level, we developed tools to reduce BCI calibration times. We notably developed algorithms to calibrate CSP filters and LDA classifiers with little training data, either by re-using data from other subjects, or by generating artificial data from the data available. Our evaluations suggested these methods could be used to calibrate a BCI with only 10 trials per class while reaching the same accuracy as standard BCI designs with 30 trials per class, effectively reducing calibration times by 3.

At the user training level, we advocated that a major cause of MI-BCIs poor efficiency and effectiveness is inappropriate user training approaches. Indeed, we shown, both theoretically - by studying the educational science literature - and practically - by training users to perform simple motor tasks with the same feedback as for BCI, that current approaches were highly suboptimal. In an effort to redefine such training approaches, and notably to later design user-specific training approaches, we studied the impact of the user profile on BCI performances. Our results highlighted the major impact of the user's spatial abilities of BCI performances, both for pure motor imagery-based BCIs, as well as for MI-BCIs with non-motor mental tasks. In general, it seems that the user's spatial abilities, attention (both state and trait) and relationship with technology all have a significant impact on BCI performances, and should thus be considered during BCI training and use. We have also explored new types of feedback, notably multi-user feedback in a gaming environment and continuous tactile feedback, which both led to increased BCI performances. We also made a first step towards designing an explanatory feedback, illustrating the promises and challenges in that direction.

Finally, at the usage level, we showed that even though current EEG-based BCIs are neither effective nor efficient, the technology used to design them can still be used with success for other applications than communication and control. First, we explored the use of EEG-based BCIs technologies for neuroergonomics, to obtain continuous, non-interrupting and objective measures of the ergonomics pro and cons of an HCI, by analyzing the EEG signals of users interacting with this HCI. We notably showed that we could estimate workload levels continuously from EEG signals, including during complex interaction tasks such as 3D object manipulation or navigation in 3D environments. We also shown that we could estimate from EEG signals how comfortable a stereoscopic visualization was. Second, we have designed 3 systems to enable anyone to visualize his/her own brain activity and mental states, or those of others, in real-time. These systems, named the Mind-Mirror, Teegi and Tobe, opened interesting application perspectives in the areas of education, neurofeedback, rehabilitation or group interactions, among others.

Altogether, this work contributed new tools to improve BCI usability, both in terms of efficiency and effectiveness. It also contributed new fundamental knowledge that will be useful to further improve BCIs in the future - notably in the area of user training, and opened new research directions, e.g., on new user training paradigms, EEG-based neuroergonomics for 3DUI or *introspectibles*. Interestingly enough, several of the tools presented in this manuscript are freely available for the community, as free and open-source pieces of code, libraries and softwares, see, e.g., https://sites.google.com/site/fabienlotte, https://github.com/potioc/tobe or our real-time BCI platform OpenViBE (http://openvibe.inria.fr/). Naturally, all this work is ongoing work, with its pros and cons, and we are far from having solved the usability issue of EEG-based BCIs. There is still a lot to do, with numerous perspectives and potential research directions to explore. We discuss this in the next sections.

Limitations and perspectives of this work

Although the works presented here did contribute knowledge and tools to improve EEG-based BCIs usability, it is naturally not exempt from limitations. The first and main limitation of all this work, and notably of Chapters 1 and 2, is that the subjects from the data sets we used for signal processing, as well as the participants from the experiments we conducted, were all healthy users. There are several potential applications of EEG-based BCIs that target healthy users, such as entertainment, sport training or passive BCIs, which means our results are still applicable in practice and potentially useful. Nonetheless, the historic application of EEG-based BCIs, and still one of its main applications, is to be used as an assistive technology for paralyzed and severely motor impaired users, e.g., patients suffering from Amyotrophic Lateral Sceloris (ALS), or Spinal Cord Injuries (SCI) [268, 6], or more recently for stroke rehabilitation [9, 42]. Unfortunately it is known that the EEG signals from those categories of end-users differ, at least in part, from those of healthy users [269, 270]. As such, it is unclear whether the methods and findings from our work could be used as such with motor-impaired users, and if so, how well would that work and generalize to that population. Some works suggested that some methods and tools initial developed and tested on healthy subjects prove useful as well for ALS and stroke patients [271, 272], which is encouraging. Nevertheless, numerous adaptations are generally necessary when working with motor-impaired end-users, and we will have to face these challenges in the near-future [273]. Actually, we are currently starting a collaboration and a pilot study with the local hospital in Bordeaux to use BCI and our BCI user training approaches for stroke rehabilitation.

Another limitation, or at least a legitimate question about all this work, is the exclusive reliance on EEG as the sole measure of brain activity. There are other sensors that could be relevant and useful as well. Moreover, EEG is a very old brain imaging method [274], with inherent limitations in terms of noise sensitivity and spatial resolution. As such, EEG sensors themselves are a clear bottleneck in the BCI pipeline. One may wonder how much of this work would still be useful if a new and better sensor than EEG were to be designed. Probably that a large part of the signal processing work would turn out mostly outdated and irrelevant, although the ideas and concepts (e.g., data generation, user-to-user transfer) may be transferred. Our work on BCI user training should still prove mostly useful and relevant as it targets more fundamental concepts, more specifically how users learn to gain control over a BCI, what makes them succeed or fail and why. Finally, our work on neuroergonomics and introspectibles would probably become even better with better sensors, enabling more accurate mental states estimations.

A last main limitation, which is very common in the BCI field for good practical reasons, is the usually small size of the participants samples used to assess algorithms or validate BCI applications or training approaches. Indeed, data are costly, and conducting experiments with real-users takes a lot of time and efforts, especially if those users should come several times, e.g., for user training. We did our best to have reasonably large populations, with experiments with between 10 to 20 users each. While this is already enough to draw some statistically valid conclusions, one could still argue that such results may not generalize to a larger population or to a different population, which would be a valid critique. 10 to 15 years ago, it was not uncommon to have BCI studies published with only 4 or 5 subjects or even with only a single one. The BCI community has made progress in this area, but we should still all strive to increase even further our population sizes.

In terms of future work, there are many directions that we could and should explore. At the signal processing level, as already mentioned, we will investigate in more details Riemmanian geometry both for EEG signal representation, processing and classification, given how promising it seems [99, 98, 124]. We would also try to better understand from a fundamental point of view, and ideally computationally model, EEG signals variability, to be able to design algorithms that are invariant across time, contexts and users. Finally, most robust algorithms available for BCI, ours included, are robust at the optimization level, i.e., they can be calibrated even with noisy training EEG data. We should work on designing classifiers and filters that can also deal with noisy and/or corrupted input features, even at the testing level, during online use.

At the BCI user training level, there is a large body of research to be conducted. Actually, this will be the area in which we will devote most of our efforts in the coming few years, and we have a couple of research projects going on to do so. In particular, regarding user training, there is a substantial lack of models and fundamental knowledge about:

- How do BCI users learn BCI control?
- Why can they or can they not acquire BCI control?
- How do they perceive and process the feedback?
- What is a correctly executed mental imagery task?
- What is(are) a BCI control skill(s)?
- How to train this(these) BCI control skill(s)?
- How do the users' skills, states and traits impact BCI learning?
- What are the interaction between the signal processing methods used and BCI user learning efficiency?
- When and how to update filters and classifiers to favor learning?

among many other open research questions. We need to conduct a number of fundamental research studies to understand all these points and answer those questions. Then we need to mathematically model the BCI user and the BCI user training process to then be able to provide optimal training tasks and feedback to each user to ensure successful BCI control skill acquisition. This would thus also require interactions between the EEG signal processing level and the user training level. Hopefully, such research should enable us to drastically improve BCI usability by making users proficient in BCI control. This should in turn make BCI ready to be used in practice outside laboratories, for actual communication and control applications.

Finally, at the usage level, we plan to continue working in neuroergonomics, to design a full-fledged neuroergonomics platform and/or software tools, that can be used in practice for HCI and UX evaluation. This will notably include trying to estimate other mental states relevant for HCI evaluation such as multimodal attention, frustration, graded error recognition or engagement, among many others. Note that being able to estimate such states in brain signals would prove useful beyond neuroergonomics, and could be useful for many passive BCI applications, to study human learning and thus BCI user training as well. Regarding instrospectibles, we will further explore this concept in education and/or well-being applications. The introspectibles puppets such as Teegi or Tobe would also potentially prove useful as richer and more motivating feedback for BCI user training.

Perspectives for EEG-based BCI

In addition to specific perspectives on the work presented in this manuscript, it would be interesting to think about broader perspectives for the BCI field in general, since all our work in about BCI. Currently, EEG-based BCIs are a rapidly developing technology, involving a large number of laboratories, and raising a substantial public interest. However, despite this enthusiasm, one may wonder whether BCIs are really that promising and useful, or if they are simply a hype - fueled by their *sci-fi-ish* side - that will soon fade away. An interesting tool to try to answer that question is Gartner's "hype cycle" [275]. This "hype cycle", that can be represented as a curve, presents the different stages that any innovative technology goes through and the visibility of this technology over time. This cycle is composed of five main steps, as illustrated in Figure 3.4:

1. The technology trigger
2. The peak of inflated expectations, where much more is expected from the technology than what it can deliver, or even that it will ever deliver
3. The trough of disillusionment, when the technology is considered to be useless, or at least much less useful than what it is or what it will become
4. The slope of enlightenment where the true potential and usefulness of the technology begin to be perceived
5. The plateau of productivity, when the technology is finally useful and used where it is relevant.

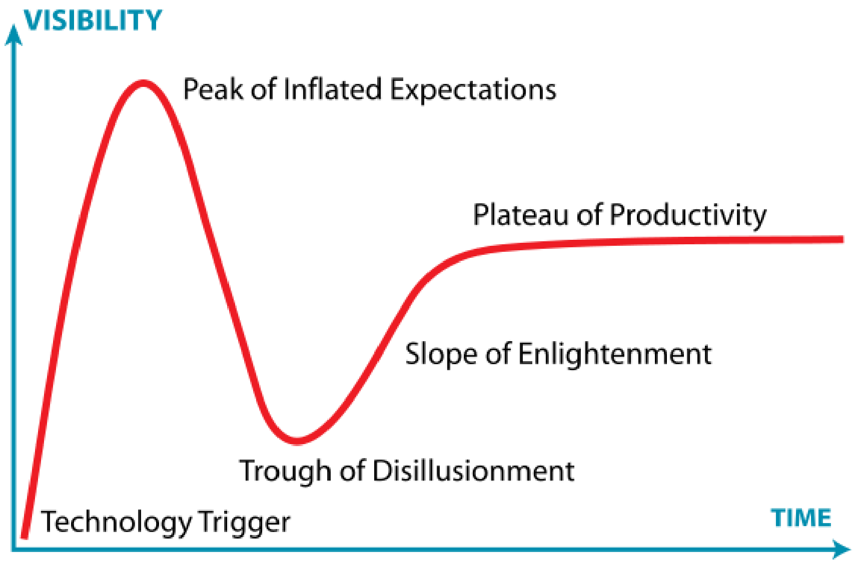


Figure 3.12: The "hype cycle": the different visibility stages of an innovative technology (source: Jeremy Kemp at English Wikipedia)

The Gartner group is at the origin of this concept and each year positions different innovative technologies on the curve. This group estimates that in 2015 (see http://www.gartner.com/newsroom/id/3114217), BCIs were located at the very beginning of this curve, that is immediately after the launch of the technology, about to climb toward the peak of exaggerated expectations. In other words, this means we should witness increased expectations about BCIs, some of which will be completely unrealistic. These unreasonable expectations are already very visible in the mainstream media, which do not hesitate to talk about "mind reading", or even "telepathy", BCI being obviously completely unable to do so. Actually, currently there are even very little evidence that they could even be able to do so in the future.

Our vision as scientists, and probably the vision of the scientific community, is certainly different from that of the general public. Indeed, we would place BCIs further on the curve, around the downward slope, moving toward the trough of disillusionment. Indeed, as already mentioned, the current performance of BCIs makes them unable to compete with most of the standard interaction peripherals, such as mice, keyboards, joysticks or eye-trackers. Thus, some researchers do not necessarily anticipate a bright future for BCIs used for communication and control [276]. Research funding agencies (notably the European ones), probably share similar fears and thus have funded two successive projects to define each time a roadmap for BCI research, and thus identify what would be the true usefulness of BCIs, beyond the exaggerated expectations [32]. From these two roadmaps (see also bnci-horizon-2020.eu/), as well as from the work of the BCI research community (see, e.g., [14]), it appears that BCI research is still a relatively young field, still offering many promising perspectives and a large number of research directions to explore. There is still much to understand about the brain, about measuring its activity and about using a BCI. Basic research on these aspects has a great potential in improve BCIs performance. Numerous practical points must also be explored 1) to make BCIs usable outside laboratories, by patients with motor disabilities, 2) to put them into the hands of the general public, and 3) to make them commercially viable and ethically usable. Finally, BCIs have until now been applied only in a relatively small number of application areas. Opening BCIs to other applications could reveal many other areas for which BCIs might be useful.

On the basic research side, it is necessary to reinforce the scientific basis of BCIs, to develop principles, theories and models that could describe BCI operation and guide BCI design. This notably includes understanding the specificities of the end-users of BCIs, the way users learn how to control BCIs (i.e., answering the research questions on BCI user training mentioned above), the reasons that make some succeed and others not, and why their performances vary in all cases. It is also necessary to understand the sources of fluctuations in brain activity, and to know how to design features and classifiers that are robust to noise and to the non-stationarity of brain signals.

We already mentioned that current BCIs were still mostly prototypes tested in laboratory, and that very few of them were really marketed and used regularly in hospitals or at home. Therefore, on the practical side, many efforts must be made in translational research, to bring BCIs out of the laboratories. This notably includes designing BCIs that are comfortable and that can be used easily and in an independent way by the patients who need them. This also includes making them cheaper and user-friendly, both for motor-impaired users and for the general public. Finally, this requires to make BCIs marketable, and to ensure that their use is ethical and beneficial to society. This thus necessitates Research and Developments (R&D) efforts at the hardware level, to design appealing, comfortable and easy-to-use sensors, at the software level, to design BCI processing pipelines that can be launched with a single click, and at the societal levels, to answer the needs of end-users and limit the risks of unethical applications.

Finally, at the application level, it is worthwhile opening BCI technologies to new applications and fields. So far, BCIs have been mostly used as communication and control systems in order to send direct commands to a machine or to a computer. However, there are many other potential applications where BCIs might be useful, or where they are already useful but are still little explored. This includes notably introducing BCIs as tools for experimental investigation in neuroscience and in psychology, for instance to design adaptive experiments [46] or to exploit machine learning tools from BCI research to perform fundamental neuroscience research, as we did to investigate speech perception and production in Electrocorticography (ECoG) signals [277]. BCI technologies can also prove useful for neuroergonomics (as discussed in length before) or adaptive man-machine interfaces, e.g., neurally informed rapid data search [278], as well as for many other applications based on passive BCI [13]. They are also promising for various forms of rehabilitation, both motor and cognitive. Many of these new alternative applications have been considered in the recent roadmap "BNCI Horizon 2020" [32]. In short, BCI technologies will be potentially useful well beyond applications aiming to directly control systems, which needs to be explored in more detail. For more discussion and perspectives on the future of BCI, the interested reader can refer to our recent book on BCI [279], and notably its perspective chapter in the second volume [280].

To conclude, despite the already impressive amount of work on BCIs that has been carried out and published, BCI research remains a very young area, in which most work and research are still to come, with very promising potential results. For a large number of emerging technologies, decades have elapsed between the first research carried out on the subject and their use in practice in the society. This is the case, for instance, of the computer mouse or the Internet. Thus, it seems reasonable to suggest that BCIs have a long life and many exciting years of research in front of them. Now is thus a great time to contribute to these R&D efforts, and we will strive to keep doing so in a relevant way.

Bibliography

[1] J.R. Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtscheller, and T.M. Vaughan. Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113(6):767–791, 2002.

[2] J Wolpaw and E.W. Wolpaw. *Brain-computer interfaces: principles and practice*. Oxford University Press, 2012.

[3] F. Lotte, L. Bougrain, and M. Clerc. Electroencephalography (eeg)-based brain-computer interfaces. In *Wiley Encyclopedia on Electrical and Electronices Engineering*. Wiley, 2015.

[4] G.E. Fabiani, D.J. McFarland, J.R. Wolpaw, and G. Pfurtscheller. Conversion of EEG activity into cursor movement by a brain-computer interface (BCI). *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 12(3):331–338, 2004.

[5] J. R. Wolpaw, D. J. McFarland, G. W. Neat, and C. A. Forneris. An EEG-based brain-computer interface for cursor control. *Electroencephalography and clinical neurophysiology*, 78:252–259, 1991.

[6] G. Pfurtscheller, G.R. Müller-Putz, R. Scherer, and C. Neuper. Rehabilitation with brain-computer interface systems. *IEEE Computer*, 41(10):58–65, 2008.

[7] H. Cecotti. Spelling with non-invasive brain-computer interfaces - current and future trends. *Journal of Physiology-Paris*, 2011.

[8] B. Rebsamen, C. Guan, H. Zhang, C. Wang, C. Teo, M.H. Ang Jr., and E. Burdet. A brain controlled wheelchair to navigate in familiar environments. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(6):590–598, 2010.

[9] Ander Ramos-Murguialday, Doris Broetz, Massimiliano Rea, Leonhard Läer, Özge Yilmaz, Fabricio L Brasil, Giulia Liberati, Marco R Curado, Eliana Garcia-Cossio, Alexandros Vyziotis, et al. Brain–machine interface in chronic stroke rehabilitation: A controlled study. *Annals of neurology*, 74(1):100–108, 2013.

[10] Kai Keng Ang and Cuntai Guan. Brain-computer interface in stroke rehabilitation. *Journal of Computing Science and Engineering*, 7(2):139–146, 2013.

[11] A. Lécuyer, F. Lotte, R.B. Reilly, R. Leeb, M. Hirose, and M. Slater. Brain-computer interfaces, virtual reality and videogames. *IEEE Computer*, 41(10):66–72, 2008.

[12] D Coyle, J Principe, F Lotte, and A Nijholt. Guest editorial: Brain/neuronal computer games interfaces and interaction. *IEEE Transactions on Computational Intelligence and AI in Games*, 5(2):77–81, 2013.

[13] T.O. Zander and C. Kothe. Towards passive brain-computer interfaces: applying brain-computer interface technology to human-machine systems in general. *Journal of Neural Engineering*, 8, 2011.

[14] J. van Erp, F. Lotte, and M. Tangermann. Brain-computer interfaces: Beyond medical applications. *IEEE Computer*, 45(4):26–34, 2012.

[15] G. Pfurtscheller and F. H. Lopes da Silva. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology*, 110(11):1842–1857, 1999.

[16] G. Pfurtscheller and C. Neuper. Motor imagery and direct brain-computer communication. *proceedings of the IEEE*, 89(7):1123–1134, 2001.

[17] Elisabeth VC Friedrich, Christa Neuper, and Reinhold Scherer. Whatever works: A systematic user-centered training protocol to optimize brain-computer interfacing individually. *PloS one*, 8(9):e76214, 2013.

[18] J.R. Millán, J. Mouri no, M. Franzé, F. Cincotti, M. Varsta, J. Heikkonen, and F. Babiloni. A local neural classifier for the recognition of eeg patterns associated to mental tasks. *IEEE transactions on neural networks*, 13(3):678–686, 2002.

[19] F.B. Vialatte, M. Maurice, J. Dauwels, and A. Cichocki. Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives. *Progress in Neurobiology,*, 90:418–438, 2010.

[20] R. Fazel-Rezai, BZ Allison, C Guger, EW Sellers, SC Kleih, and A Kübler. P300 brain computer interface: current challenges and emerging trends. *Frontiers in Neuroengineering*, 5(14), 2012.

[21] J.R. Wolpaw, G.E. Loeb, B.Z. Allison, E. Donchin, O.F. do Nascimento, W.J. Heetderks, F. Nijboer, W.G. Shain, and J. N. Turner. BCI meeting 2005–workshop on signals and recording methods. *IEEE Transaction on Neural Systems and Rehabilitation Engineering.*, 14(2):138–141, 2006.

[22] J. Mellinger, G. Schalk, C. Braun, H. Preissl, W. Rosenstiel, N. Birbaumer, and A. Kübler. An MEG-based brain-computer interface (BCI). *Neuroimage*, 36(3):581–593, 2007.

[23] N. Weiskopf, K. Mathiak, S.W. Bock, F. Scharnowski, R. Veit, W. Grodd, R. Goebel, and N. Birbaumer. Principles of a brain-computer

interface (BCI) based on real-time functional magnetic resonance imaging (fMRI). *IEEE Transactions on Biomedical Engineering*, 51(6):966–970, 2004.

[24] M.A. Lebedev and M.A.L. Nicolelis. Brain-machine interfaces: past, present and future. *Trends in Neurosciences*, 29(9):536–546, 2006.

[25] E. Niedermeyer and F. Lopes da Silva. *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, ISBN 0781751268, 5th edition, 2005.

[26] A. Bashashati, M. Fatourechi, R. K. Ward, and G. E. Birch. A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals. *Journal of Neural engineering*, 4(2):R35–57, 2007.

[27] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi. A review of classification algorithms for EEG-based brain-computer interfaces. *Journal of Neural Engineering*, 4:R1–R13, 2007.

[28] C. Neuper and G. Pfurtscheller. *Brain-Computer Interfaces*, chapter Neurofeedback Training for BCI Control, pages 65–78. The Frontiers Collection, 2010.

[29] G. Pfurtscheller, D. Flotzinger, and J. Kalcher. Brain-computer interface-a new communication device for handicapped persons. *journal of microcomputer application*, 16:293–299, 1993.

[30] N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kübler, J. Perelmouter, E. Taub, and H. Flor. A spelling device for the paralysed. *Nature*, 398:297–298, 1999.

[31] L.A. Farwell and E. Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology*, 70:510–523, 1988.

[32] Clemens Brunner, Niels Birbaumer, Benjamin Blankertz, Christoph Guger, Andrea Kübler, Donatella Mattia, José del R Millán, Felip Miralles, Anton Nijholt, Eloy Opisso, et al. Bnci horizon 2020: towards a roadmap for the bci community. *Brain-Computer Interfaces*, (ahead-of-print):1–10, 2015.

[33] Bernhard Graimann, Brendan Allison, and Gert Pfurtscheller. *Brain-computer interfaces: Revolutionizing human-computer interaction*. Springer, 2010.

[34] Gert Pfurtscheller, Brendan Z Allison, Günther Bauernfeind, Clemens Brunner, Teodoro Solis Escalante, Reinhold Scherer, Thorsten O Zander, Gernot Mueller-Putz, Christa Neuper, and Niels Birbaumer. The hybrid BCI. *Frontiers in neuroscience*, 4:3, 2010.

[35] Gernot Muller-Putz, Robert Leeb, Michael Tangermann, Johannes Hohne, Andrea Kubler, Febo Cincotti, Donatella Mattia, Rudiger Rupp, Klaus-Robert Muller, and J Del R Millan. Towards non-invasive hybrid brain-computer interfaces: framework, practice, clinical application, and beyond. *Proceedings of the IEEE*, 103(6):926–943, 2015.

[36] R. Sitaram, H. Zhang, C. Guan, M. Thulasidas, Y. Hoshi, A. Ishikawa, K. Shimizu, and N. Birbaumer. Temporal classification of multi-channel near infrared spectroscopy signals of motor imagery for developing a brain-computer interface. *NeuroImage*, 34(4):1416–1427, 2007.

[37] Audrey Girouard, Erin Treacy Solovey, Leanne M Hirshfield, Evan M Peck, Krysta Chauncey, Angelo Sassaroli, Sergio Fantini, and Robert JK Jacob. From brain signals to adaptive interfaces: using fnirs in hci. In *Brain-Computer Interfaces*, pages 221–237. Springer, 2010.

[38] Siamac Fazli, Jan Mehnert, Jens Steinbrink, Gabriel Curio, Arno Villringer, Klaus-Robert Müller, and Benjamin Blankertz. Enhanced performance by a hybrid NIRS–EEG brain computer interface. *Neuroimage*, 59(1):519–529, 2012.

[39] G. Schalk, K.J. Miller, N.R. Anderson, J.A. Wilson, M.D. Smyth, J.G. Ojemann, D.W. Moran, J.R. Wolpaw, and E.C. Leuthardt. Two-dimensional movement control using electrocorticographic signals in humans. *Journal of Neural Engineering*, 5(1):75–84, 2008.

[40] Scott Makeig, Christian Kothe, Tim Mullen, Nima Bigdely-Shamlo, Zhilin Zhang, and Kenneth Kreutz-Delgado. Evolving signal processing for brain–computer interfaces. *Proceedings of the IEEE*, 100(Special Centennial Issue):1567–1584, 2012.

[41] Clemens Brunner, Giuseppe Andreoni, Lugi Bianchi, Benjamin Blankertz, Christian Breitwieser, Shinichiro Kanoh, Christian A Kothe, Anatole Lécuyer, Scott Makeig, Jürgen Mellinger, et al. BCI software platforms. In *Towards Practical Brain-Computer Interfaces*, pages 303–331. Springer, 2013.

[42] Kai Keng Ang and Cuntai Guan. Brain–computer interface for neurorehabilitation of upper limb after stroke. *Proceedings of the IEEE*, 103(6):944–953, 2015.

[43] C. Mühl, B. Allison, A. Nijholt, and G. Chanel. A survey of affective brain computer interfaces: principles, state-of-the-art, and challenges. *Brain-Computer Interfaces*, pages 1–19, 2014.

[44] F. Lotte, J. Fujisawa, H. Touyama, R. Ito, M. Hirose, and A. Lécuyer. Towards ambulatory brain-computer interfaces: A pilot study with P300 signals. In *5th Advances in Computer Entertainment Technology Conference (ACE)*, pages 336–339, 2009.

[45] Cornelia Kranczioch, Catharina Zich, Irina Schierholz, and Annette Sterr. Mobile EEG and its potential to promote the theory and application of imagery-based motor rehabilitation. *International Journal of Psychophysiology*, 91(1):10–15, 2014.

[46] Gaëtan Sanchez, Jean Daunizeau, Emmanuel Maby, Olivier Bertrand, Aline Bompas, and Jérémie Mattout. Toward a new application of real-time electrophysiology: online optimization of cognitive neurosciences hypothesis testing. *Brain sciences*, 4(1):49–72, 2014.

[47] C. Guger, G. Edlinger, W. Harkam, I. Niedermayer, and G. Pfurtscheller. How many people are able to operate an EEG-based brain-computer interface (BCI)? *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2):145–147, 2003.

[48] C. Guger, S. Daban, E. Sellers, C. Holzner, G. Krausz, R. Carabalona, F. Gramatica, and G. Edlinger. How many people are able to control a P300-based brain-computer interface (BCI)? *Neuroscience Letters*, 462(1):94–98, 2009.

[49] Christoph Guger, Brendan Z Allison, Bernhard Großwindhager, Robert Prückl, Christoph Hintermüller, Christoph Kapeller, Markus Bruckner, Gunther Krausz, and Günter Edlinger. How many people could use an ssvep bci? *Frontiers in neuroscience*, 6, 2012.

[50] B Blankertz, C Sannelli, S Halder, EM Hammer, A Kübler, K-R Müller, G Curio, and T Dickhaus. Neurophysiological predictor of SMR-based BCI performance. *NeuroImage*, 51(4):1303–1309, 2010.

[51] N Neumann and N Birbaumer. Predictors of successful self control during brain-computer communication. *Journal of Neurology, Neurosurgery & Psychiatry*, 74:1117–21, 2003.

[52] A Kübler and N Birbaumer. Brain–computer interfaces and communication in paralysis: extinction of goal directed thinking in completely paralysed patients? *Clinical neurophysiology*, 119(11):2658–2666, 2008.

[53] B.Z. Allison and C. Neuper. Could anyone use a BCI? In Desney S. Tan and Anton Nijholt, editors, *Brain-Computer Interfaces*, volume 0 of *Human-Computer Interaction Series*, pages 35–54. Springer London, 2010.

[54] Stephanie Brandl, Johannes Hohne, Klaus-Robert Muller, and Wojciech Samek. Bringing BCI into everyday life: Motor imagery in a pseudo realistic environment. In *Neural Engineering (NER), 2015 7th International IEEE/EMBS Conference on*, pages 224–227. IEEE, 2015.

[55] M. Fatourechi, A. Bashashati, R. Ward, and G. Birch. EMG and EOG artifacts in brain computer interface systems: A survey. *Clinical Neurophysiology*, 118(3):480–494, 2007.

[56] DJ Krusienski, M Grosse-Wentrup, F Galán, D Coyle, KJ Miller, E Forney, and CW Anderson. Critical issues in state-of-the-art brain-computer interface signal processing. *Journal of Neural Engineering*, 8(2):025002, 2011.

[57] P. Shenoy, M. Krauledat, B. Blankertz, R.P.N. Rao, and K.-R. Müller. Towards adaptive classification for bci. *Journal of Neural Engineering*, 3(1):R13, 2006.

[58] Christian Mühl, Camille Jeunet, and Fabien Lotte. EEG-based workload estimation across affective contexts. *Frontiers in Neuroscience section Neuroprosthetics*, 8:114, 2014.

[59] Fabien Lotte. Signal processing approaches to minimize or suppress calibration time in oscillatory activity-based brain–computer interfaces. *Proceedings of the IEEE*, 2015.

[60] Fabien Lotte and Camille Jeunet. Towards improved BCI based on human learning principles. In *3rd International Brain-Computer Interfaces Winter Conference*, 2015.

[61] Jeffrey Rubin and Dana Chisnell. *Handbook of usability testing: how to plan, design and conduct effective tests*. John Wiley & Sons, 2008.

[62] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R. Müller. Optimizing spatial filters for robust EEG single-trial analysis. *IEEE Signal Proc Magazine*, 25(1):41–56, 2008.

[63] D. J. McFarland, C. W. Anderson, K.-R. Müller, A. Schlögl, and D. J. Krusienski. BCI meeting 2005-workshop on BCI signal processing: feature extraction and translation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(2):135 – 138, 2006.

[64] The Future BNCI consortium. Future BNCI: A roadmap for future directions in brain / neuronal computer interaction. Technical report, 2012.

[65] N. Brodu, F. Lotte, and A. Lécuyer. Exploring two novel features for EEG-based brain-computer interfaces: Multifractal cumulants and predictive complexity. *Neurocomputing*, 79(1):87–94, 2012.

[66] F. Lotte. A new feature and associated optimal spatial filter for EEG signal classification: Waveform length. In *International Conference on Pattern Recognition (ICPR)*, pages 1302–1305, 2012.

[67] N Caramia, F Lotte, and S Ramat. Optimizing spatial filter pairs for eeg classification based on phase synchronization. In *International Conference on Audio, Speech and Signal Processing (ICASSP'2014)*, 2014.

[68] H. Wendt, P. Abry, and S. Jaffard. Bootstrap for empirical multifractal analysis. *IEEE Signal Processing Magazine*, (38), 2007.

[69] J.-F. Muzy, E. Bacry, and A. Arneodo. Multifractal formalism for fractal signals: The structure-function approach versus the wavelet-transform modulus-maxima method. *Physical Review E.*, 47:875–884, 1993.

[70] P. Abry, S. Jaffard, and B. Lashermes. Revisiting scaling, multifractal, and multiplicative cascades with the wavelet leader lens. In *SPIE*, pages 103–117, 2004.

[71] B. Mandelbrot, A. Fisher, and L. Calvet. A multifractal model of asset returns. Technical Report 1164, Cowles Fundation, 1997.

[72] Nicolas Brodu. Reconstruction of epsilon-machines in predictive frameworks and decisional states. *Advances in Complex Systems*, 14(05):761–794, 2011.

[73] C.R. Shalizi. *Causal Architecture, Complexity, and Self-Organization in Time Series and Cellular Automata*. PhD thesis, University of Wisconsin at Madison, 2001.

[74] J. Crutchfield and K. Young. Inferring statistical complexity. *Physical Review Letters*, 62(2):105–108, 1989.

[75] D. Tkach, H. Huang, and T. Kuiken. Study of stability of time-domain features for electromyographic pattern recognition. *Journal of NeuroEngineering and Rehabilitation*, 7(1):21, 2010.

[76] Jean-Philippe Lachaux, Eugenio Rodriguez, Jacques Martinerie, Francisco J Varela, et al. Measuring phase synchrony in brain signals. *Human brain mapping*, 8(4):194–208, 1999.

[77] Jean-Philippe Lachaux, Eugenio Rodriguez, Michel Le van Quyen, Antoine Lutz, Jacques Martinerie, and Francisco J Varela. Studying single-trials of phase synchronous activity in the brain. *International Journal of Bifurcation and Chaos*, 10(10):2429–2439, 2000.

[78] Qingguo Wei, Yijun Wang, Xiaorong Gao, and Shangkai Gao. Amplitude and phase coupling measures for feature extraction in an eeg-based brain–computer interface. *Journal of Neural Engineering*, 4(2):120, 2007.

[79] DJ Krusienski, DJ McFarland, and JR Wolpaw. Value of amplitude, phase, and coherence features for a sensorimotor rhythm-based brain-computer interface. *Brain research bulletin*, 87(1):130–134, 2012.

[80] B. Blankertz, K. R. Müller, G. Curio, T. M. Vaughan, G. Schalk, J. R. Wolpaw, A. Schlögl, C. Neuper, G. Pfurtscheller, T. Hinterberger, M. Schröder, and N. Birbaumer. The BCI competition 2003: Progress and perspectives in detection and discrimination of EEG single trials. *IEEE Transactions on Biomedical Engineering*, 51(6):1044–1051, 2004.

[81] S. Lemm, C. Schafer, and G. Curio. BCI competition 2003–data set III: probabilistic modeling of sensorimotor mu rhythms for classification of imaginary hand movements. *IEEE Transactions on Biomedical Engineering*, 51(6):1077–1080, 2004.

[82] B. Blankertz, K. R. Müller, D. J. Krusienski, G. Schalk, J. R. Wolpaw, A. Schlögl, G. Pfurtscheller, J.R. Míllan, M. Schröder, and N. Birbaumer. The BCI competition III: Validating alternative approaches to actual BCI problems. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(2):153–159, 2006.

[83] M Tangermann, KR Müller, A Aertsen, N Birbaumer, C Braun, C Brunner, R Leeb, C Mehring, K J Miller, GR Müller-Putz, G Nolte, G Pfurtscheller, H Preissl, G Schalk, A Schlögl, C Vidaurre, S Waldert, and B Blankertz. Review of the bci competition iv. *Frontiers in Neuroscience*, 6, 2012.

[84] M. Naeem, C. Brunner, R. Leeb, B. Graimann, and G. Pfurtscheller. Seperability of four-class motor imagery data using independent components analysis. *Journal of Neural Engineering*, 3:208–216, 2006.

[85] A Schlögl, C Keinrath, D Zimmermann, R Scherer, R Leeb, and G Pfurtscheller. A fully automated correction method of eog artifacts in eeg recordings. *Clinical Neurophysiology*, 118(1):98–104, 2007.

[86] M. Grosse-Wentrup and M. Buss. Multi-class common spatial pattern and information theoretic feature extraction. *IEEE Transactions on Biomedical Engineering*, 55(8):1991–2000, 2008.

[87] M. Grosse-Wentrup, C. Liefhold, K. Gramann, and M. Buss. Beamforming in non invasive brain computer interfaces. *IEEE Transactions on Biomedical Engineering*, 56(4):1209 – 1219, 2009.

[88] B. Reuderink and M. Poel. Robustness of the common spatial patterns algorithm in the BCI-pipeline. Technical report, HMI, University of Twente, 2008.

[89] F. Lotte and C.T. Guan. Regularizing common spatial patterns to improve BCI designs: Unified theory and new algorithms. *IEEE Transactions on Biomedical Engineering*, 58(2):355–362, 2011.

[90] O. Ledoit and M. Wolf. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2):365–411, 2004.

[91] G. Dornhege, B. Blankertz, G. Curio, and K.R. Müller. Boosting bit rates in non-invasive EEG single-trial classifications by feature combination and multi-class paradigms. *IEEE Transactions on Biomedical Engineering*, 51(6):993–1002, 2004.

[92] A. Schlögl, F. Lee, H. Bischof, and G. Pfurtscheller. Characterization of four-class motor imagery EEG data for the BCI-competition 2005. *Journal of Neural Engineering*, pages L14–L22, 2005.

[93] B. Blankertz, M. Kawanabe, R. Tomioka, F. Hohlefeld, V. Nikulin, and K.-R. Müller. Invariant common spatial patterns: Alleviating nonsta-

tionarities in brain-computer interfacing. In *Advances in Neural Information Processing Systems 20, In . MIT Press, Cambridge, MA*, 2008.

[94] W. Samek, C. Vidaurre, K.-R. Müller, and M. Kawanabe. Stationary common spatial patterns for brain-computer interfacing. *Journal of Neural Engineering*, 9(2), 2012.

[95] W Samek, M Kawanabe, and K Muller. Divergence-based framework for common spatial patterns algorithms. *IEEE Reviews in Biomedical Engineering*, 2014.

[96] J. Farquhar, N.J. Hill, T.N. Lal, and B. Schölkopf. Regularised CSP for sensor selection in BCI. In *Proceedings of the 3rd international BCI workshop*, 2006.

[97] M. Arvaneh, C. Guan, K.K. Ang, and H.C. Quek. Optimizing the channel selection and classification accuracy in eeg-based bci. *IEEE Transactions on Biomedical Engineering*, 58:1865–1873, 2011.

[98] Alexandre Barachant, Stéphane Bonnet, Marco Congedo, and Christian Jutten. Multiclass brain–computer interface classification by riemannian geometry. *Biomedical Engineering, IEEE Transactions on*, 59(4):920–928, 2012.

[99] Marco Congedo, Alexandre Barachant, and Anton Andreev. A new generation of brain-computer interface based on riemannian geometry. *arXiv preprint arXiv:1310.8115*, 2013.

[100] Florian Yger, Fabien Lotte, and Masashi Sugiyama. Averaging covariance matrices for eeg signal classification based on the csp: an empirical study. In *Proc. EUSIPCO*, pages 2721–2725, 2015.

[101] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications*, 2007.

[102] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten. Classification of covariance matrices using a Riemannian-based kernel for BCI applications. *Neurocomputing*, 2013.

[103] R. Bhatia. *Positive Definite Matrices.* Princeton University Press, 2009.

[104] M. Moakher. A differential geometric approach to the geometric mean of symmetric positive-definite matrices. *SIAM J. Mat. An. App.*, 2005.

[105] B. Jeuris, R. Vandebril, and B. Vandereycken. A survey and comparison of contemporary algorithms for computing the matrix geometric mean. *Elec. Trans. Num. An.*, 2012.

[106] D. Bini and B. Iannazzo. Computing the Karcher mean of symmetric positive definite matrices. *Linear Algebra and its Applications*, 2013.

[107] T. Fletcher, S. Venkatasubramanian, and S. Joshi. The geometric median on Riemannian manifolds with application to robust atlas estimation. *NeuroImage*, 2009.

[108] Inbal Horev, Florian Yger, and Masashi Sugiyama. Geometry-aware principal component analysis for symmetric positive definite matrices. In *JMLR: Workshop and Conference Proceedings*, 2015.

[109] M. Kawanabe, C. Vidaurre, S. Schoeller, B. Blankertz, and K.R. Mueller. Robust common spatial filters with a maxmin approach. In *EMBS-Conference*, 2009.

[110] B. Blankertz, G. Dornhege, M. Krauledat, G. Curio, and K.-R. Müller. The non-invasive Berlin brain-computer interface: Fast acquisition of effective performance in untrained subjects. *NeuroImage*, 37(2):539–550, 2007.

[111] B. Blankertz, F. Losch, M. Krauledat, G. Dornhege, G. Curio, and K.-R. Müller. The berlin brain-computer interface: Accurate performance from first-session in BCI-naive subjects. *IEEE Trans Biomed Eng*, 55(10):2452–2462, 2008.

[112] F. Lotte and C.T. Guan. Learning from other subjects helps reducing brain-computer interface calibration time. In *Proc. ICASSP*, pages 614–617, 2010.

[113] H. Lu, H.-L. Eng, C. Guan, K.N. Plataniotis, and A.N. Venetsanopoulos. Regularized common spatial pattern with aggregation for EEG classification in small-sample setting. *IEEE Transactions on Biomedical Engineering*, 57(12):2936–2946, 2010.

[114] H. Kang, Y. Nam, and S. Choi. Composite common spatial pattern for subject-to-subject transfer. *IEEE Signal Processing Letters*, 16(8):683 – 686, 2009.

[115] R. Chakraborty and U. Garain. Role of synthetically generated samples on speech recognition in a resource-scarce language. *Proc. ICPR*, pages 1618–1621, 2010.

[116] H. Mouchère, S. Bayoudh, E. Anquetil, and L. Miclet. Synthetic on-line handwriting generation by distortions and analogy. In *Proc. IGS*, pages 10–13, 2007.

[117] F. Lotte, C. Guan, and K.K. Ang. Comparison of designs towards a subject-independent brain-computer interface based on motor imagery. In *Proc. EMBC*, pages 4543 – 4546, 2009.

[118] Siamac Fazli, Florin Popescu, Márton Danóczy, Benjamin Blankertz, Klaus-Robert Müller, and Cristian Grozea. Subject-independent mental state classification in single trials. *Neural networks*, 22(9):1305–1312, 2009.

[119] C. Jeunet, F. Lotte, and C. Mühl. Design and validation of a mental and social stress induction protocol: Towards load-invariant physiology-based detection. In *International conference on Physiological Computing Systems (PhyCS 2014)*, pages 98–106, 2014.

[120] Camille Jeunet, Bernard N 'kaoua, Sriram Subramanian, Martin Hachet, and Fabien Lotte. Predicting Mental Imagery-Based BCI Performance from Personality, Cognitive Profile and Neurophysiological Patterns. *PLoS ONE*, page 20, July 2015.

[121] J. Schumacher, C. Jeunet, and F. Lotte. Towards explanatory feedback for user training in brain-computer interfaces. In *IEEE SMC*, pages 3169–3174, 2015.

[122] Emmanuel Kalunga, Sylvain Chevallier, and Quentin Barthélemy. Data augmentation in riemannian space for brain-computer interfaces. In *ICML Workshop on Statistics, Machine Learning and Neuroscience (Stamlins)*, 2015.

[123] Florian Yger and Masashi Sugiyama. Supervised logeuclidean metric learning for symmetric positive definite matrices. *arXiv preprint arXiv:1502.03505*, 2015.

[124] F. Yger, M. Berar, and F. Lotte. Riemannian approaches in brain-computer interfaces: a review. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2016.

[125] Jeremiah D Wander, Timothy Blakely, Kai J Miller, Kurt E Weaver, Lise A Johnson, Jared D Olson, Eberhard E Fetz, Rajesh PN Rao, and Jeffrey G Ojemann. Distributed cortical adaptation during learning of a brain–computer interface task. *Proceedings of the National Academy of Sciences*, 110(26):10818–10823, 2013.

[126] B. Blankertz, G. Dornhege, M. Krauledat, K.-R. Müller, V. Kunzmann, F. Losch, and G. Curio. The Berlin Brain-Computer Interface: EEG-based communication without subject training. *IEEE Trans. Neural Sys. Rehab. Eng.*, 14(2):147–152, 2006.

[127] F Lotte, F Larrue, and C Mühl. Flaws in current human training protocols for spontaneous brain-computer interfaces: lessons learned from instructional design. *Frontiers in Human Neuroscience*, 7(568), 2013.

[128] MD Merrill. First principles of instruction: a synthesis. *Trends and issues in instructional design and technology*, 2:62–71, 2007.

[129] J Hattie and H Timperley. The power of feedback. *Review of Educational Research*, 77:81–112, 2007.

[130] V.J. Shute. Focus on formative feedback. *Review of Educational Research*, 78:153–189, 2008.

[131] J Sweller, JJG van Merrienboer, and FGWC Pass. Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3):251–296, 1998.

[132] Roxana Moreno. Decreasing cognitive load for novice students: Effects of explanatory versus corrective feedback in discovery-

based multimedia. *Instructional Science*, 32:99–113, 2004. 10.1023/B:TRUC.0000021811.66966.1d.

[133] S Ainsworth. DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16:183–198, 2006.

[134] Gerald R Salancik and Jeffrey Pfeffer. A social information processing approach to job attitudes and task design. *Administrative science quarterly*, pages 224–253, 1978.

[135] David W Johnson and Roger T Johnson. An educational psychology success story: Social interdependence theory and cooperative learning. *Educational researcher*, 38(5):365–379, 2009.

[136] Sonja C Kleih and Andrea Kübler. Psychological factors influencing brain-computer interface (BCI) performance. In *Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on*, pages 3192–3196. IEEE, 2015.

[137] C. Jeunet, E. Jahanpour, and F. Lotte. Why standard brain-computer interface (BCI) training protocols should be changed: An experimental study. *Journal of Neural Engineering*, 13(3):036024, 2016.

[138] G. Müller-Putz, R. Scherer, C. Brunner, R. Leeb, and G. Pfurtscheller. Better than random: a closer look on BCI results. *International Journal of Bioelectromagnetism*, 10(1):52–55, 2008.

[139] Eva Maria Hammer, Sebastian Halder, Benjamin Blankertz, Claudia Sannelli, Thorsten Dickhaus, Sonja Kleih, Klaus-Robert Müller, and Andrea Kübler. Psychological predictors of smr-bci performance. *Biological psychology*, 89(1):80–86, 2012.

[140] Minkyu Ahn and Sung Chan Jun. Performance variation in motor imagery brain–computer interface: A brief review. *Journal of neuroscience methods*, 243:103–110, 2015.

[141] A.B. Randolph. Not all created equal: Individual-technology fit of brain-computer interfaces. In *HICSS*, pages 572–578, 2012.

[142] Adriane B Randolph, Melody Moore Jackson, and Saurav Karmakar. Individual characteristics and their effect on predicting mu rhythm modulation. *Intl. Journal of Human–Computer Interaction*, 27(1):24–37, 2010.

[143] E.-M. Hammer, T. Kaufmann, S. C. Kleih, B. Blankertz, and A. Kübler. Visuo-motor coordination ability predicts performance with brain-computer interfaces controlled by modulation of sensorimotor rhythms (SMR). *Frontiers in Human Neuroscience*, 8, 2014.

[144] David Wechsler. Wechsler adult intelligence scale–fourth edition (wais–iv). *San Antonio, TX: NCS Pearson*, 2008.

[145] Daniel B Berch, Robert Krikorian, and Eileen M Huha. The corsi block-tapping task: Methodological and theoretical considerations. *Brain and cognition*, 38:317–38, 1998.

[146] Arthur Lester Benton. *The revised visual retention test: clinical and experimental applications.* State University of Iowa; Distributor for the USA, Psychological Corporation, New York, 1963.

[147] David A Kolb. *Learning style inventory.* McBer and Company, 1999.

[148] Raymond B Cattell and Heather EP Cattell. Personality structure and the new fifth edition of the 16pf. *Educational and Psychological Measurement*, 55:926–37, 1995.

[149] Hanna Levenson. Activism and powerful others: Distinctions within the concept of internal-external control. *Journal of personality assessment*, 38:377–83, 1974.

[150] C D Spielberger, R L Gorsuch, and R E Lushene. *Manual for the State-Trait Anxiety Inventory.* Consulting Psychologists Press, Palo Alto, 1970.

[151] Robert H Bruininks. *Bruininks-Oseretsky test of motor proficiency: Examiner's manual.* American Guidance Service Circle Pines, MN, 1978.

[152] S. G. Vandenberg and A. R. Kuse. Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and motor skills*, 47:599–604, 1978.

[153] W.S. Noble. How does multiple testing correction work? *Nature Biotechnology*, 27:1135–37, 2009.

[154] Shelley Derksen and HJ Keselman. Backward, forward and stepwise automated subset selection algorithms: Frequency of obtaining authentic and noise variables. *British Journal of Mathematical and Statistical Psychology*, 45:265–82, 1992.

[155] Mark J Whittingham, Philip A Stephens, Richard B Bradbury, and Robert P Freckleton. Why do we still use stepwise modelling in ecology and behaviour? *Journal of animal ecology*, 75:1182–89, 2006.

[156] Steven E Poltrock and Polly Brown. Individual differences in visual imagery and spatial ability. *Intelligence*, 8:93–138, 1984.

[157] Byron P Rourke and M Alan J Finlayson. Neuropsychological significance of variations in patterns of academic performance: Verbal and visual-spatial abilities. *Journal of Abnormal Child Psychology*, 6:121–33, 1978.

[158] Anke Vromen, Jeanine A Verbunt, Sascha Rasquin, and Derick T Wade. Motor imagery in patients with a right hemisphere stroke and unilateral neglect. *Brain Injury*, 25:387–93, 2011.

[159] Noriko Hara. Student distress in a web-based distance education course. *Information, Communication & Society*, 3:557–79, 2001.

[160] M. G. Moore. Learner autonomy: The second dimension of independant lerning. *Convergence*, 5:76–88, 1972.

[161] Nicholas LeBoutillier and David F Marks. Mental imagery and creativity: A meta-analytic review study. *British Journal of Psychology*, 94:29–44, 2003.

[162] R. Scherer, J. Faller, D. Balderas, E.V.C. Friedrich, M. Pröll, B. Allison, and G. Müller-Putz. Brain–computer interfacing: more than the sum of its parts. *Soft Computing*, 17(2):317–331, 2013.

[163] Camille Jeunet. Training users' spatial abilities to improve brain-computer interface performance: A theoretical approach. In *9th Conference of Young Researchers in Cognitives Sciences*, 2015.

[164] William F Chaplin, Oliver P John, and Lewis R Goldberg. Conceptions of states and traits: dimensional attributes with ideals as prototypes. *Journal of personality and social psychology*, 54(4):541, 1988.

[165] F Nijboer, A Furdea, I Gunst, J Mellinger, DJ MacFarland, N Birbaumer, and A Kübler. An auditory brain-computer interface. *Journal of Neuroscience Methods*, 167(1):43–50, 2007.

[166] Mechthildis Regina Averbeck. *Skalen zur Erfassung der Lebensqualität: SEL*. Swets Test Services, 1997.

[167] Irene Daum, Brigitte Rockstroh, Niels Birbaumer, Thomas Elbert, Anthony Canavan, and Werner Lutzenberger. Behavioural treatment of slow cortical potentials in intractable epilepsy: neuropsychological predictors of outcome. *Journal of Neurology, Neurosurgery & Psychiatry*, 56(1):94–97, 1993.

[168] M. Grosse-Wentrup, B. Schölkopf, and J. Hill. Causal influence of gamma oscillations on the sensorimotor rhythm. *NeuroImage*, pages 1–26, 2011.

[169] Moritz Grosse-Wentrup and Bernhard Schölkopf. High gamma-power predicts performance in sensorimotor-rhythm brain–computer interfaces. *Journal of neural engineering*, 9(4):046001, 2012.

[170] SC Kleih, T Kaufmann, E Hammer, I Pisotta, F Pichiorri, A Riccio, D Mattia, and A Kübler. Motivation and smr-bci: fear of failure affects bci performance. In *Proceedings of the Fifth International Brain-Computer Interface Meeting 2013*. Verlag der Technischen Universität Graz Austria, 2013.

[171] Wenke Burde and Benjamin Blankertz. *Is the locus of control of reinforcement a predictor of brain-computer interface performance?* na, 2006.

[172] Femke Nijboer, Niels Birbaumer, and Andrea Kübler. The influence of psychological state and motivation on brain–computer interface performance in patients with amyotrophic lateral sclerosis–a longitudinal study. *Frontiers in neuroscience*, 4:55, 2010.

[173] Matthias Witte, Silvia Erika Kober, Manuel Ninaus, Christa Neuper, and Guilherme Wood. Control beliefs can predict the ability to up-regulate sensorimotor rhythm during neurofeedback training. *Frontiers in human neuroscience*, 2013.

[174] Aleksandra Vuckovic and Bethel A Osuagwu. Using a motor imagery questionnaire to estimate the performance of a brain–computer interface based on object oriented motor imagery. *Clinical Neurophysiology*, 124(8):1586–1595, 2013.

[175] S. Teillet, F. Lotte, B. N'Kaoua, and C. Jeunet. Towards a spatial ability training to improve mental imagery based brain-computer interface (MI-BCI) performance: a pilot study. In *Proc. IEEE SMC*, 2016.

[176] L. Bonnet, F. Lotte, and A. Lécuyer. Two brains, one game: Design and evaluation of a multi-user BCI video game based on motor imagery. *IEEE Transactions on Computational Intelligence and AI in Games*, 5(2):185–198, 2013.

[177] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Transactions on Rehabilitation Engineering*, 8(4):441–446, 2000.

[178] Kiuk Gwak, Robert Leeb, José del R Millán, and Dae-Shik Kim. Quantification and reduction of visual load during bci operation. In *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 2795–2800. IEEE, 2014.

[179] Robert Leeb, Kiuk Gwak, Dae-Shik Kim, Jose del Millan, et al. Freeing the visual channel by exploiting vibrotactile bci feedback. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, pages 3093–3096. IEEE, 2013.

[180] Daniel Afergan, Tomoki Shibata, Samuel W Hincks, Evan M Peck, Beste F Yuksel, Remco Chang, and Robert JK Jacob. Brain-based target expansion. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*, pages 583–593. ACM, 2014.

[181] Sang-Hoon Han and Min-Shik Kim. Visual search does not remain efficient when executive working memory is working. *Psychological Science*, 15(9):623–628, 2004.

[182] C. Jeunet, C. Vi, D. Spelmezan, B. N'Kaoua, F. Lotte, and S. Subramanian. Continuous tactile feedback for motor-imagery based brain-computer interaction in a multitasking context. In *Proc. Interact 2015*, 2015.

[183] Marieke E Thurlings, Jan BF van Erp, Anne-Marie Brouwer, Benjamin Blankertz, and Peter Werkhoven. Control-display mapping in brain–computer interfaces. *Ergonomics*, 55(5):564–580, 2012.

[184] George A Gescheider, John H Wright, and Ronald T Verrillo. *Information-processing channels in the tactile sensory system: A psychophysical and physiological analysis*. Psychology Press, 2010.

[185] I.I Goncharova, D.J McFarland, T.M Vaughan, and J.R Wolpaw. Emg contamination of eeg: spectral and topographical characteristics. *Clinical Neurophysiology*, 114(9):1580 – 1593, 2003.

[186] Moritz Grosse-Wentrup and Bernhard Schölkopf. A brain–computer interface based on self-regulation of gamma-oscillations in the superior parietal cortex. *J. neur. eng.*, 11(5):056015, 2014.

[187] Laurent George, Laurent Bonnet, and Anatole Lécuyer. Freeze the BCI until the user is ready: a pilot study of a bci inhibitor. In *International BCI conference*, 2011.

[188] Camille Jeunet, Bernard N'Kaoua, and Fabien Lotte. Advances in user-training for mental-imagery-based bci control: Psychological and cognitive factors and their neural correlates. *Progress in brain research*, 2016.

[189] Roger N'Kambou, Jacqueline Bourdeau, and Riichiro Mizoguchi. *Advances in intelligent tutoring systems*. Springer, 2010.

[190] Camille Jeunet, Bernard N'Kaoua, Roger N'Kambou, and Fabien Lotte. Why and how to use intelligent tutoring systems to adapt mi-bci training to each user. In *6th International BCI Meeting*, 2016.

[191] E Thomas, M Dyson, and M Clerc. An analysis of performance evaluation for motor-imagery based bci. *Journal of neural engineering*, 10(3):031001, 2013.

[192] M. Gomez Rodriguez, J. Peters, J. Hill, B. Schölkopf, A. Gharabaghi, and M. Grosse-Wentrup. Closing the sensorimotor loop: haptic feedback helps decoding of motor imagery. *Journal of Neural Engineering*, 2011.

[193] Fabien Lotte, Josef Faller, Christoph Guger, Yann Renard, Gert Pfurtscheller, Anatole Lécuyer, and Robert Leeb. Combining BCI with Virtual Reality: Towards new applications and improved BCI. In Brendan Z. Allison, Stephen Dunne, Robert Leeb, José Del R. Millán, and Anton Nijholt, editors, *Towards Practical Brain-Computer Interfaces*, Biological and Medical Physics, Biomedical Engineering, pages 197–220. Springer Berlin Heidelberg, 2013.

[194] Reinhold Scherer, Andreas Schwarz, Gernot R Müller-Putz, Viktoria Pammer-Schindler, and Mariano Lloria Garcia. Game-based bci training: Interactive design for individuals with cerebral palsy. In *Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on*, pages 3175–3180. IEEE, 2015.

[195] R. Parasuraman. Neuroergonomics: research and practice. *Theoretical Issues in Ergonomic Science*, 4(1-2):5–20, 2003.

[196] R. Parasuraman and G.F. Wilson. Putting the brain to work: Neuroergonomics past, present, and future. *Human Factors*, 50(3):468–474, 2008.

[197] Mickael Causse, Frédéric Dehais, Patrice Péran, Umberto Sabatini, and Josette Pastor. The effects of emotion on pilot decision-making: A neuroergonomic approach to aviation safety. *Transportation research part C: emerging technologies*, 33:272–281, 2013.

[198] Effie Lai-Chong Law, Virpi Roto, Marc Hassenzahl, Arnold POS Vermeeren, and Joke Kort. Understanding, scoping and defining user experience: a survey approach. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 719–728. ACM, 2009.

[199] Arnold POS Vermeeren, Effie Lai-Chong Law, Virpi Roto, Marianna Obrist, Jettie Hoonhout, and Kaisa Väänänen-Vainio-Mattila. User experience evaluation methods: current state and development needs. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*, pages 521–530. ACM, 2010.

[200] Jérémy Frey. *Leveraging human-computer interactions and social presence with physiological computing.* PhD thesis, Université de Bordeaux, 2016.

[201] Michel Beaudouin-Lafon. Designing interaction, not interfaces. In *Proceedings of the working conference on Advanced visual interfaces*, pages 15–22. ACM, 2004.

[202] Brad A Myers. A brief history of human-computer interaction technology. *interactions*, 5(2):44–54, 1998.

[203] D.A. Bowman, E. Kruijff, J.J. LaViola Jr., and I. Poupyrev. *3D User Interfaces: Theory and Practice.* Addison-Wesley/Pearson Education, 2005.

[204] Jacek Jankowski, Martin Hachet, et al. A survey of interaction techniques for interactive 3d environments. In *Eurographics 2013-STAR*, 2013.

[205] D. Bowman, J. Gabbard, and D. Hix. A survey of usability evaluation in virtual environments: Classification and comparison of methods. *Presence*, 11, 2002.

[206] Jérémy Frey, Christian Mühl, Fabien Lotte, and Martin Hachet. Review of the use of electroencephalography as an evaluation method for human-computer interaction. In *Proc. of PhyCS*, 2014.

[207] SG Hart and LE Staveland. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Human mental workload*. 1988.

[208] Richard E. Nisbett and Timothy DeCamp Wilson. Telling more than we can know: Verbal reports on mental processes. *Psych. Rev.*, 84(3):231–260, 1977.

[209] Rosalind W. Picard. Affective computing. Technical Report 321, MIT Media Laboratory, 1995.

[210] J. Matias Kivikangas, Inger Ekman, Guillaume Chanel, Simo Järvelä, Ben Cowley, Pentti Henttonen, and Niklas Ravaja. Review on psychophysiological methods in game research. *Proc. of 1st Nordic DiGRA*, 2010.

[211] C Berka, DJ Levendowski, MN Lumicao, A Yau, G Davis, VT Zivkovic, RE Olmstead, PD Tremoulet, and PL Craven. Eeg correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, space, and environmental medicine*, 78(Supplement 1):B231–B244, 2007.

[212] Benjamin Blankertz, Michael Tangermann, Carmen Vidaurre, Siamac Fazli, Claudia Sannelli, Stefan Haufe, Cecilia Maeder, Lenny Ramsey, Irene Sturm, Gabriel Curio, and Klaus-Robert Müller. The Berlin Brain-Computer Interface: Non-Medical Uses of BCI Technology. *Front Neurosci*, 4(December):198, January 2010.

[213] Anne-Marie Brouwer, Maarten a Hogervorst, Jan B F van Erp, Tobias Heffelaar, Patrick H Zimmerman, and Robert Oostenveld. Estimating workload using EEG spectral power and ERPs in the n-back task. *J. of neur. engin.*, 9(4):045008, August 2012.

[214] David Grimes, DS Tan, and SE Hudson. Feasibility and pragmatics of classifying working memory load with an electroencephalograph. *CHI '08*, page 835, 2008.

[215] Laurent George and A Lécuyer. An overview of research on'passive'brain-computer interfaces for implicit human-computer interaction. In *ICABB 2010*, 2010.

[216] W Klimesch, M Doppelmayr, H Russegger, T Pachinger, and J Schwaiger. Induced alpha band power changes in the human EEG and attention. *Neuroscience letters*, 244(2):73–6, March 1998.

[217] Romain E Trachel, Maureen Clerc, and Thomas G Brochier. Decoding covert shifts of attention induced by ambiguous visuospatial cues. *Frontiers in human neuroscience*, 9, 2015.

[218] François Laurent, Mario Valderrama, Michel Besserve, Mathias Guillard, Jean-Philippe Lachaux, Jacques Martinerie, and Geneviève Florence. Multimodal information improves the rapid detection of mental fatigue. *Biomed. Sig. Proc. Contr.*, pages 1–9, March 2013.

[219] Sylvie Charbonnier, Raphaelle N Roy, Stephane Bonnet, and Aurelie Campagne. EEG index for control operators' mental fatigue monitoring

using interactions between brain regions. *Expert Systems with Applications*, 52:91–98, 2016.

[220] P.W. Ferrez and J.R. Millán. Error-related EEG potentials generated during simulated brain-computer interaction. *IEEE Transactions on Biomedical Engineering*, 55(3):923–929, 2008.

[221] Ricardo Chavarriaga, Aleksander Sobolewski, and José del R Millán. Errare machinale est: the use of error-related potentials in brain-machine interfaces. *Frontiers in neuroscience*, 8, 2014.

[222] Chi Vi and Sriram Subramanian. Detecting error-related negativity for interaction design. *CHI '12*, page 493, 2012.

[223] Lennart E. Nacke, S. Stellmach, and Craig A. Lindley. Electroencephalographic Assessment of Player Experience: A Pilot Study in Affective Ludology. *SAG*, 42(5):632–655, August 2010.

[224] Riccardo Berta, Francesco Bellotti, Alessandro De Gloria, Danu Pranantha, and Carlotta Schatten. Electroencephalogram and Physiological Signal Analysis for Assessing Flow in Games. *IEEE Trans. Comp. Intel. and AI in Games*, 5(2):164–175, June 2013.

[225] Abhijit Karnik, Tovi Grossman, and Sriram Subramanian. Comparison of User Performance in Mixed 2D-3D Multi-Display Environments. In *INTERACT '13*, Lecture Notes in Computer Science, 2013.

[226] Evan M M Peck, Beste F Yuksel, Alvitta Ottley, Robert JK Jacob, and Remco Chang. Using fnirs brain sensing to evaluate information visualization interfaces. In *Proc. CHI*, pages 473–482, 2013.

[227] Erik W Anderson, Kristin C Potter, Laura E Matzen, Jason F Shepherd, Gilbert A Preston, and Cláudio T Silva. A user study of visualization effectiveness using eeg and cognitive load. In *Eurographics*, pages 791–800, 2011.

[228] S Mathan, T Feyereisen, and S Whitlow. WorkSense: Exploring the Feasibility of Human Factors Assessment Using Electrophysiological Sensors. In *Proc. ICACS*, Santa Monica, 2007.

[229] C Kirschbaum, K-M Pirke, and D H Hellhammer. The "Trier Social Stress Test": a Tool for Investigating Psychobiological Stress Responses in a Laboratory Setting. *Neuropsychobiology*, 28:76–81, 1993.

[230] Sally S Dickerson and Margaret E Kemeny. Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychol. Bull.*, 130(3):355–391, 2004.

[231] R E Krout. Music listening to facilitate relaxation and promote wellness: Integrated aspects of our neurophysiological response to music. *The Arts in Psychotherapy*, 34:134–141, 2007.

[232] W K Kirchner. Age differences in short-term retention of rapidly changing information. *J. Exp. Psychol.*, 55(4):352–358, 1958.

[233] KK Ang, ZY Chin, C Wang, C Guan, and H Zhang. Filter bank common spatial pattern algorithm on bci competition iv datasets 2a and 2b. *Frontiers in Neuroscience*, 6, 2012.

[234] H. Peng, F. Long, and C. Ding. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1226–1238, 2005.

[235] Dennis Wobrock, Jérémy Frey, Delphine Graef, Jean-Baptiste de la Rivière, Julien Castet, and Fabien Lotte. Continuous mental effort evaluation during 3d object manipulation tasks based on brain and physiological signals. In *Proceedings of INTERACT'15*, 2015.

[236] W Samek, F C Meinecke, and K-R Muller. Transferring subspaces between subjects in brain–computer interfacing. *Biomedical Engineering, IEEE Transactions on*, 60(8):2289–2298, 2013.

[237] Fabien Lotte. A tutorial on EEG signal-processing techniques for mental-state recognition in brain–computer interfaces. In *Guide to Brain-Computer Music Interfacing*, pages 133–161. Springer, 2014.

[238] Jean-Baptiste de la Rivière, Cédric Kervégant, Emmanuel Orvain, and Nicolas Dittlo. CubTile: a multi-touch cubic interface. In *Proc. ACM VRST*, pages 69–72, 2008.

[239] Jérémy Frey, Maxime Daniel, Martin Hachet, Julien Castet, and Fabien Lotte. Framework for electroencephalography-based evaluation of user experience. In *Procedings of CHI*, pages 2283–2294, 2016.

[240] S Sternberg. High-speed scanning in human memory. *Science (New York, N.Y.)*, 153:652–654, 1966.

[241] G. Michael Poor, Guy Zimmerman, Dale S. Klopfer, Samuel D. Jaffee, Laura Marie Leventhal, and Julie Barnes. Mobility Matters: Identifying Cognitive Demands That Are Sensitive to Orientation. In *Interact '13*, volume 8117 LNCS, pages 193–210, 2013.

[242] J. Frey, A. Appriou, F. Lotte, and M. Hachet. Classifying EEG signals during stereoscopic visualization to estimate visual comfort. *Computational Intelligence & Neuroscience*, 2016.

[243] Giannis Drossis, Dimitris Grammenos, Ilia Adami, and Constantine Stephanidis. 3D Visualization and Multimodal Interaction with Temporal Information Using Timelines. In *INTERACT '13*, 2013.

[244] Marc Lambooij, Wijnand IJsselsteijn, Marten Fortuin, and Ingrid Heynderickx. Visual Discomfort and Visual Fatigue of Stereoscopic Displays: A Review. *J Imaging Sci Technol*, 53(3):030201, 2009.

[245] Takashi Shibata, Joohwan Kim, David M Hoffman, and Martin S Banks. The zone of comfort: Predicting visual discomfort with stereo displays. *J. Vis.*, 11, 2011.

[246] A Delorme and S Makeig. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Meth.*, 2004.

[247] U. Hoffmann, J. Vesin, and T. Ebrahimi. Spatial filters for the classification of event-related potentials. In *European Symposium on Artificial Neural Networks (ESANN 2006)*, 2006.

[248] B. Blankertz, S. Lemm, M.S. Treder, S. Haufe, and K.-R. Müller. Single-trial analysis and classification of ERP components - a tutorial. *Neuroimage*, 2010.

[249] J. Frey, Hachet, and F. Lotte. EEG-based neuroergonomics for 3D user interfaces: opportunities and challenges. *Le Travail Humain*, 2016.

[250] Gautier Durantin, J-F Gagnon, Sébastien Tremblay, and Frédéric Dehais. Using near infrared spectroscopy and heart rate variability to detect mental overload. *Behavioural brain research*, 259:16–23, 2014.

[251] Gautier Durantin, Sébastien Scannella, Thibault Gateau, Arnaud Delorme, and Frédéric Dehais. Processing functional near infrared spectroscopy signal with a kalman filter to assess working memory during simulated flight. *Frontiers in human neuroscience*, 9, 2015.

[252] Thibault Gateau, Gautier Durantin, Francois Lancelot, Sebastien Scannella, and Frederic Dehais. Real-time state estimation in a flight simulator using fnirs. *PloS one*, 10(3):e0121279, 2015.

[253] Jonathan Mercier-Ganady, Fabien Lotte, Emilie Loup-Escande, Maud Marchal, and Anatole Lecuyer. The mind-mirror: See your brain in action in your head using eeg and augmented reality. In *Virtual Reality (VR), 2014 iEEE*, pages 33–38. IEEE, 2014.

[254] Emmanuel Dubois, Philip Gray, and Laurence Nigay. *The engineering of mixed reality systems*. Springer Science & Business Media, 2009.

[255] Jérémy Frey, Renaud Gervais, Stéphanie Fleck, Fabien Lotte, and Martin Hachet. Teegi: tangible eeg interface. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*, pages 301–308. ACM, 2014.

[256] S. Vosniadou, C. Ioannides, A. Dimitrakopoulou, and E. Papademetriou. Designing learning environments to promote conceptual change in science. *Learning and Instruction*, (11):381–419, 2001.

[257] Stéphanie Fleck and Gilles Simon. An augmented reality environment for astronomy learning in elementary grades: An exploratory study. In *25ème conférence francophone sur l'Interaction Homme-Machine-IHM 2013*, page 9.

[258] Eva Hornecker. The role of physicality in tangible and embodied interactions. *Interactions*, 18(2):19, 2011.

[259] Renaud Gervais, Jérémy Frey, Alexis Gay, Fabien Lotte, and Martin Hachet. Tobe: Tangible out-of-body experience. In *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*, pages 227–235. ACM, 2016.

[260] Ramesh Raskar and G Welch. Shader lamps: Animating real objects with image-based illumination. In *Eurographics*, 2001.

[261] Michael Schmitz. *Tangible Interaction with Anthropomorphic Smart Objects in Instrumented Environments*. PhD thesis, 2010.

[262] Rollin McCraty, Mike Atkinson, Dana Tomasino, and Raymond Trevor Bradley. The coherent heart: Heart-brain interactions, psychophysiological coherence, and the emergence of system-wide order. *Integral Review*, 5(2):10–115, 2009.

[263] Suzana Herculano-Houzel. Do you know your brain? a survey on public neuroscience literacy at the closing of the decade of the brain. *The Neuroscientist*, 8(2):98–110, 2002.

[264] Daniel J Simons and Christopher F Chabris. What people believe about how memory works: A representative survey of the us population. *PloS one*, 6(8):e22757, 2011.

[265] Elliott Hedman, Lucy Miller, Sarah Schoen, and Darci Nielsen. Measuring Autonomic Arousal During Therapy. In *D&E '12*, pages 11–14, 2012.

[266] Mirko Raca and Pierre Dillenbourg. System for Assessing Classroom Attention. In *LAK '13*, pages 1–5, 2013.

[267] Renaud Gervais, Joan Sol Roo, Jérémy Frey, and Martin Hachet. Introspectibles: Tangible Interaction to Foster Introspection. In *Computing and Mental Health - CHI '16 Workshop*, San Jose, United States, May 2016.

[268] J.R. Millán, R. Rupp, G. Müller-Putz, R. Murray-Smith, C. Giugliemma, M. Tangermann, C. Vidaurre, F. Cincotti, A. Kübler, R. Leeb, C. Neuper, K.-R. Müller, and D. Mattia. Combining brain-computer interfaces and assistive technologies: State-of-the-art and challenges. *Frontiers in Neuroprosthetics*, 2010.

[269] N. Neumann and A. Kubler. Training locked-in patients: a challenge for the use of brain-computer interfaces. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2):169–172, 2003.

[270] Niels Birbaumer, Ander Ramos Murguialday, Moritz Wildgruber, and Leonardo G Cohen. Brain-computer-interface (bci) in paralysis. In *The European Image of God and Man*, pages 483–492. Brill, 2010.

[271] A. Kübler, F. Nijboer, J. Mellinger, T.M. Vaughan, H. Pawelzik, G. Schalk, D.J. McFarland, N. Birbaumer, and J.R. Wolpaw. Patients with ALS can use sensorimotor rhythms to operate a brain-computer interface. *Neurology*, 34:1775–1777, 2005.

[272] K.K. Ang, C. Guan, K.S.G. Chua, and B.T. Ang. A large clinical study on the ability of stroke patients in using eeg-based motor imagery brain-computer interface. *Clinical EEG and Neuroscience*, 2011.

[273] Andrea Kübler, Donatella Mattia, Rüdiger Rupp, and Michael Tangermann. Facing the challenge: Bringing brain–computer interfaces to end-users. *Artificial intelligence in medicine*, 59(2):55–60, 2013.

[274] H. Berger. Ueber das elektroenkephalogramm des menschen. *Archiv für Psychiatrie und Nervenkrankheiten*, 87:527–570, 1929.

[275] Alexander Linden and Jackie Fenn. Understanding gartner's hype cycles. *Strategic Analysis Report Nº R-20-1971. Gartner, Inc*, 2003.

[276] Femke Nijboer, Jens Clausen, Brendan Z Allison, and Pim Haselager. The asilomar survey: Stakeholders' opinions on ethical issues related to brain-computer interfacing. *Neuroethics*, 6(3):541–578, 2013.

[277] Fabien Lotte, Jonathan Brumberg, Peter Brunner, Aysegul Gunduz, Anthony L. Ritaccio, Cuntai Guan, and Gerwin Schalk. Electrocorticographic Representations of Segmental Features in Continuous Speech. *Frontiers in Human Neuroscience*, 9(97):27, February 2015.

[278] Adam D Gerson, Lucas C Parra, and Paul Sajda. Cortically coupled computer vision for rapid image search. *IEEE Transactions on neural systems and rehabilitation engineering*, 14(2):174–179, 2006.

[279] M. Clerc, L. Bougrain, and F. Lotte. *Brain-Computer Interfaces 1: Foundations and Methods*. ISTE-Wiley, 2016.

[280] M. Clerc, L. Bougrain, and F. Lotte. *Brain-Computer Interfaces 2: Technology and Applications*. ISTE-Wiley, 2016.